Graduate Texts in Mathematics

Melvyn B. Nathanson

Elementary Methods in Number Theory

Springer

Graduate Texts in Mathematics 195

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Graduate Texts in Mathematics

(continued after index)

Melvyn B. Nathanson

Elementary Methods in Number Theory

Springer

Melvyn B. Nathanson
Department of Mathematics
Lehman College (CUNY)
Bronx, NY 10468
USA
nathansn@alpha.lehman.cuny.edu

Mathematics Subject Classification (1991): 11-01

Library of Congress Cataloging-in-Publication Data
Nathanson, Melvyn B. (Melvyn Bernard), 1944 –
Elementary methods in number theory / Melvyn B. Nathanson.
p. cm.—(Graduate texts in mathematics; 195)
Includes bibliographical references and index.
ISBN 0-387-98912-9 (hardcover: alk. paper)
1. Number theory. I. Title. II. Series.
QA241.N3475 2000
512′.7—dc21 99-42812

ISBN 0-387-98912-9 Springer-Verlag New York Berlin Heidelberg SPIN 10742484

To Paul Erdős,

1913–1996,

a friend and collaborator for 25 years, and a
master of elementary methods in number theory.

Preface

Arithmetic is where numbers run across your mind looking for
the answer.
Arithmetic is like numbers spinning in your head faster and
faster until you blow up with the answer.
KABOOM!!!
Then you sit back down and begin the next problem.

Alexander Nathanson [99]

This book, *Elementary Methods in Number Theory*, is divided into three parts.

Part I, "A first course in number theory," is a basic introduction to elementary number theory for undergraduate and graduate students with no previous knowledge of the subject. The only prerequisites are a little calculus and algebra, and the imagination and perseverance to follow a mathematical argument. The main topics are divisibility and congruences. We prove Gauss's law of quadratic reciprocity, and we determine the moduli for which primitive roots exist. There is an introduction to Fourier analysis on finite abelian groups, with applications to Gauss sums. A chapter is devoted to the *abc conjecture*, a simply stated but profound assertion about the relationship between the additive and multiplicative properties of integers that is a major unsolved problem in number theory.

The "first course" contains all of the results in number theory that are needed to understand the author's graduate texts, *Additive Number Theory: The Classical Bases* [104] and *Additive Number Theory: Inverse Problems and the Geometry of Sumsets* [103].

The second and third parts of this book are more difficult than the "first course," and require an undergraduate course in advanced calculus or real analysis.

Part II is concerned with prime numbers, divisors, and other topics in multiplicative number theory. After deriving properties of the basic arithmetic functions, we obtain important results about divisor functions, and we prove the classical theorems of Chebyshev and Mertens on the distribution of prime numbers. Finally, we give elementary proofs of two of the most famous results in mathematics, the *prime number theorem*, which states that the number of primes up to x is asymptotically equal to $x/\log x$, and *Dirichlet's theorem* on the infinitude of primes in arithmetic progressions.

Part III, "Three problems in additive number theory," is an introduction to some classical problems about the additive structure of the integers. The first additive problem is *Waring's problem*, the statement that, for every integer $k \geq 2$, every nonnegative integer can be represented as the sum of a bounded number of kth powers. More generally, let $f(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0$ be an integer-valued polynomial with $a_k > 0$ such that the integers in the set $A(f) = \{f(x) : x = 0, 1, 2, \ldots\}$ have no common divisor greater than one. Waring's problem for polynomials states that every sufficiently large integer can be represented as the sum of a bounded number of elements of $A(f)$.

The second additive problem is *sums of squares*. For every $s \geq 1$ we denote by $R_s(n)$ the number of representations of the integer n as a sum of s squares, that is, the number of solutions of the equation

$$n = x_1^2 + \cdots + x_s^2$$

in integers $x_1, \ldots, x_s$. The shape of the function $R_s(n)$ depends on the parity of s. In this book we derive formulae for $R_s(n)$ for certain even values of s, in particular, for $s = 2, 4, 6, 8$, and 10.

The third additive problem is the *asymptotics of partition functions*. A partition of a positive integer n is a representation of n in the form $n = a_1 + \cdots + a_k$, where the parts $a_1, \ldots, a_k$ are positive integers and $a_1 \geq \cdots \geq a_k$. The partition function $p(n)$ counts the number of partitions of n. More generally, if A is any nonempty set of positive integers, the partition function $p_A(n)$ counts the number of partitions of n with parts belonging to the set A. We shall determine the asymptotic growth of $p(n)$ and, more generally, of $p_A(n)$ for any set A of integers of positive density.

This book contains many examples and exercises. By design, some of the exercises require old-fashioned manipulations and computations with pencil and paper. A few exercises require a calculator. Number theory, after all, begins with the positive integers, and students should get to know and love them.

This book is also an introduction to the subject of "elementary methods in analytic number theory." The theorems in this book are simple statements about integers, but the standard proofs require contour integration,

modular functions, estimates of exponential sums, and other tools of complex analysis. This is not unfair. In mathematics, when we want to prove a theorem, we may use any method. The rule is “no holds barred.” It is OK to use complex variables, algebraic geometry, cohomology theory, and the kitchen sink to obtain a proof. But once a theorem is proved, once we know that it is true, particularly if it is a simply stated and easily understood fact about the natural numbers, then we may want to find another proof, one that uses only “elementary arguments” from number theory. Elementary proofs are not better than other proofs, nor are they necessarily easy. Indeed, they are often technically difficult, but they do satisfy the aesthetic boundary condition that they use only arithmetic arguments.

This book contains elementary proofs of some deep results in number theory. We give the Erdős-Selberg proof of the prime number theorem, Linnik’s solution of Waring’s problem, Liouville’s still mysterious method to obtain explicit formulae for the number of representations of an integer as the sum of an even number of squares, and Erdős’s method to obtain asymptotic estimates for partition functions. Some of these proofs have not previously appeared in a text. Indeed, many results in this book are new.

Number theory is an ancient subject, but we still cannot answer the simplest and most natural questions about the integers. Important, easily stated, but still unsolved problems appear throughout the book. You should think about them and try to solve them.

Melvyn B. Nathanson[1]
Maplewood, New Jersey
November 1, 1999

[1]Supported in part by grants from the PSC-CUNY Research Award Program and the NSA Mathematical Sciences Program. This book was completed while I was visiting the Institute for Advanced Study in Princeton, and I thank the Institute for its hospitality. I also thank Jacob Sturm for many helpful discussions about parts of this book.

Notation and Conventions

We denote the set of positive integers (also called the natural numbers) by $\mathbf{N}$ and the set of nonnegative integers by $\mathbf{N}_0$. The integer, rational, real, and complex numbers are denoted by $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$, respectively. The absolute value of $z \in \mathbf{C}$ is $|z|$. We denote by $\mathbf{Z}^n$ the group of lattice points in the n-dimensional Euclidean space $\mathbf{R}^n$.

The integer part of the real number x, denoted by $[x]$, is the largest integer that is less than or equal to x. The fractional part of x is denoted by $\{x\}$. Then $x = [x] + \{x\}$, where $[x] \in \mathbf{Z}$, $\{x\} \in \mathbf{R}$, and $0 \leq \{x\} < 1$. In computer science, the integer part of x is often called the *floor* of x, and denoted by $\lfloor x \rfloor$. The smallest integer that is greater than or equal to x is called the *ceiling* of x and denoted by $\lceil x \rceil$.

We adopt the standard convention that an empty sum of numbers is equal to 0 and an empty product is equal to 1. Similarly, an empty union of subsets of a set X is equal to the empty set, and an empty intersection is equal to X.

We denote the *cardinality* of the set X by $|X|$. The largest element in a finite set of numbers is denoted by $\max(X)$ and the smallest is denoted by $\min(X)$.

Let a and d be integers. We write $d|a$ if d divides a, that is, if there exists an integer q such that $a = dq$. The integers a and b are called *congruent* modulo m, denoted by $a \equiv b \pmod{m}$, if m divides $a - b$.

A *prime number* is an integer $p > 1$ whose only divisors are 1 and p. The set of prime numbers is denoted by $\mathbf{P}$, and p_k is the kth prime. Thus, $p_1 = 2, p_2 = 3, \ldots, p_{11} = 31, \ldots$. Let p be a prime number. We write $p^r \| n$

if p^r is the largest power of p that divides the integer n, that is, p^r divides n but p^{r+1} does not divide n.

The *greatest common divisor* and the *least common multiple* of the integers $a_1, \ldots, a_k$ are denoted by $(a_1, \ldots, a_k)$ and $[a_1, \ldots, a_k]$, respectively. If A is a nonempty set of integers, then $\gcd(A)$ denotes the greatest common divisor of the elements of A.

The *principle of mathematical induction* states that if $S(k)$ is some statement about integers $k \geq k_0$ such that $S(k_0)$ is true and such that the truth of $S(k-1)$ implies the truth of $S(k)$, then $S(k)$ holds for all integers $k \geq k_0$. This is equivalent to the *minimum principle*: A nonempty set of integers bounded below contains a smallest element.

Let f be a complex-valued function with domain D, and let g be a function on D such that $g(x) > 0$ for all $x \in D$. We write $f \ll g$ or $f = O(g)$ if there exists a constant $c > 0$ such that $|f(x)| \leq cg(x)$ for all $x \in D$. Similarly, we write $f \gg g$ if there exists a constant $c > 0$ such that $|f(x)| \geq cg(x)$ for all $x \in D$. For example, $f \gg 1$ means that $f(x)$ is uniformly bounded away from 0, that is, there exists a constant $c > 0$ such that $|f(x)| \geq c$ for all $x \in D$. We write $f \ll_{k,\ell,\ldots} g$ if there exists a positive constant c that depends on the variables $k, \ell, \ldots$ such that $|f(x)| \leq cg(x)$ for all $x \in D$. We define $f \gg_{k,\ell,\ldots} g$ similarly. The functions f and g are called *asymptotic* as x approaches a if $\lim_{x\to a} f(x)/g(x) = 1$. Positive-valued functions f and g with domain D have the same *order of magnitude* if $f \ll g \ll f$, or equivalently, if there exist positive constants c_1 and c_2 such that $c_1 \leq f(x)/g(x) \leq c_2$ for all $x \in D$. The *counting function* of a set A of integers counts the number of positive integers in A that do not exceed x, that is,

$$A(x) = \sum_{\substack{a \in A \\ 1 \leq a \leq x}} 1.$$

Using the counting function, we can associate various densities to the set A. The *Shnirel'man density* of A is

$$\sigma(A) = \inf_{n \to \infty} \frac{A(n)}{n}.$$

The *lower asymptotic density* of A is

$$d_L(A) = \liminf_{n \to \infty} \frac{A(n)}{n}.$$

The *upper asymptotic density* of A is

$$d_U(A) = \limsup_{n \to \infty} \frac{A(n)}{n}.$$

If $d_L(A) = d_U(A)$, then $d(A) = d_L(A)$ is called the *asymptotic density* of A, and

$$d(A) = \lim_{n \to \infty} \frac{A(n)}{n}.$$

Let A and B be nonempty sets of integers and $d \in \mathbf{Z}$. We define the *sumset*

$$A + B = \{a + b : a \in A, b \in B\},$$

the *difference set*

$$A - B = \{a - b : a \in A, b \in B\},$$

the *product set*

$$AB = \{ab : a \in A, b \in B\},$$

and the *dilation*

$$d * A = \{d\}A = \{da : a \in A\}.$$

The sets A and B *eventually coincide*, denoted by $A \sim B$, if there exists an integer n_0 such that $n \in A$ if and only if $n \in B$ for all $n \geq n_0$.

We use the following arithmetic functions:

$v_p(n)$	the exponent of the highest power of p that divides n
$\varphi(n)$	Euler phi function
$\mu(n)$	Möbius function
$d(n)$	the number of divisors of n
$\sigma(n)$	the sum of the divisors of n
$\pi(x)$	the number of primes not exceeding x
$\vartheta(x), \psi(x)$	Chebyshev's functions
$\ell(n)$	$\log n$ if n is prime and 0 otherwise
$\omega(n)$	the number of distinct prime divisors of n
$\Omega(n)$	the total number of prime divisors of n
$L(n)$	$\log n$, the natural logarithm of n
$\Lambda(n)$	von Mangoldt function
$\Lambda_2(n)$	generalized von Mangoldt function
$1(n)$	1 for all n
$\delta(n)$	1 if $n = 1$ and 0 if $n \geq 2$

A *ring* is always a ring with identity. We denote by $R^\times$ the multiplicative group of units of R. A commutative ring R is a field if and only if $R^\times = R \setminus \{0\}$. If $f(t)$ is a polynomial with coefficients in the ring R, then $N_0(f)$ denotes the number of distinct zeros of $f(t)$ in R. We denote by $M_n(R)$ the ring of $n \times n$ matrices with coefficients in R.

In the study of Liouville's method, we use the symbol

$$\{f(\ell)\}_{n=\ell^2} = \begin{cases} 0 & \text{if } n \text{ is not a square,} \\ f(\ell) & \text{if } n = \ell^2,\ \ell \geq 0. \end{cases}$$

Contents

Part I

A First Course in Number Theory

1

Divisibility and Primes

1.1 Division Algorithm

Divisibility is a fundamental concept in number theory. Let a and d be integers. We say that d is a *divisor* of a, and that a is a *multiple* of d, if there exists an integer q such that

$$a = dq.$$

If d divides a, we write

$$d|a.$$

For example, 1001 is divisible by 7 and 13. Divisibility is transitive: If a divides b and b divides c, then a divides c (Exercise 14).

The *minimum principle* states that every nonempty set of integers bounded below contains a smallest element. For example, a nonempty set of nonnegative integers must contain a smallest element. We can see the necessity of the condition that the nonempty set be bounded below by considering the example of the set $\mathbf{Z}$ of all integers, positive, negative, and zero.

The minimum principle is all we need to prove the following important result.

Theorem 1.1 (Division algorithm) *Let a and d be integers with $d \geq 1$. There exist unique integers q and r such that*

$$a = dq + r \tag{1.1}$$

and

$$0 \leq r \leq d - 1. \tag{1.2}$$

The integer q is called the *quotient* and the integer r is called the *remainder* in the division of a by d.

Proof. Consider the set S of nonnegative integers of the form

$$a - dx$$

with $x \in \mathbf{Z}$. If $a \geq 0$, then $a = a - d \cdot 0 \in S$. If $a < 0$, let $x = -y$, where y is a positive integer. Since d is positive, we have $a - dx = a + dy \in S$ if y is sufficiently large. Therefore, S is a nonempty set of nonnegative integers. By the minimum principle, S contains a smallest element r, and $r = a - dq \geq 0$ for some $q \in \mathbf{Z}$. If $r \geq d$, then

$$0 \leq r - d = a - d(q+1) < r$$

and $r - d \in S$, which contradicts the minimality of r. Therefore, q and r satisfy conditions (1.1) and (1.2).

Let q_1, r_1, q_2, r_2 be integers such that

$$a = dq_1 + r_1 = dq_2 + r_2 \qquad \text{and} \qquad 0 \leq r_1, r_2 \leq d - 1.$$

Then

$$|r_1 - r_2| \leq d - 1$$

and

$$d(q_1 - q_2) = r_2 - r_1.$$

If $q_1 \neq q_2$, then

$$|q_1 - q_2| \geq 1$$

and

$$d \leq d|q_1 - q_2| = |r_2 - r_1| \leq d - 1,$$

which is impossible. Therefore, $q_1 = q_2$ and $r_1 = r_2$. This proves that the quotient and remainder are unique. □

For example, division of 16 by 7 gives the quotient 2 and the remainder 2, that is,

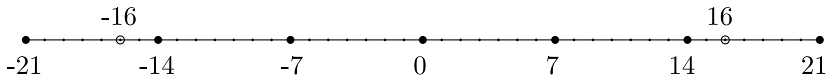
$$16 = 7 \cdot 2 + 2.$$

Division of -16 by 7 gives the quotient -3 and the remainder 5, that is,

$$-16 = 7(-3) + 5.$$

A simple geometric way to picture the division algorithm is to imagine the real number line with dots at the positive integers. Let q be a positive integer, and put a large dot on each multiple of q. The integer a either lies on one of these large dots, in which case a is a multiple of q, or a lies on a dot strictly between two large dots, that is, between two successive

multiples of q, and the distance r between a and the largest multiple of q that is less than a is a positive integer no greater than $q-1$. For example, if $q=7$ and $a=\pm 16$, we have the following picture.



The *principle of mathematical induction* states that if $S(k)$ is some statement about integers $k \geq k_0$ such that $S(k_0)$ is true and such that the truth of $S(k-1)$ implies the truth of $S(k)$, then $S(k)$ holds for all integers $k \geq k_0$. Another form of the principle of mathematical induction states that if $S(k_0)$ is true and if the truth of $S(k_0), S(k_0+1), \ldots, S(k-1)$ implies the truth of $S(k)$, then $S(k)$ holds for all integers $k \geq k_0$. Mathematical induction is equivalent to the minimum principle (Exercise 18).

Using mathematical induction and the division algorithm, we can prove the existence and uniqueness of m-adic representations of integers.

Theorem 1.2 *Let m be an integer, $m \geq 2$. Every positive integer n can be represented uniquely in the form*

$$n = a_0 + a_1 m + a_2 m^2 + \cdots + a_k m^k, \tag{1.3}$$

where k is the nonnegative integer such that

$$m^k \leq n < m^{k+1}$$

and $a_0, a_1, \ldots, a_k$ are integers such that

$$1 \leq a_k \leq m-1$$

and

$$0 \leq a_i \leq m-1 \qquad \text{for } i = 0, 1, 2, \ldots, k-1.$$

This is called the *m-adic representation* of n. The integers a_i are called the *digits* of n to *base m*. Equivalently, we can write

$$n = \sum_{i=0}^{\infty} a_i m^i,$$

where $0 \leq a_i \leq m-1$ for all i, and $a_i = 0$ for all sufficiently large integers i.

Proof. For $k \geq 0$, let $S(k)$ be the statement that every integer in the interval $m^k \leq n < m^{k+1}$ has a unique m-adic representation. We use induction on k. The statement $S(0)$ is true because if $1 \leq n < m$, then $n = a_0$ is the unique m-adic representation.

Let $k \geq 1$, and assume that the statements $S(0), S(1), \ldots, S(k-1)$ are true. We shall prove $S(k)$. Let $m^k \leq n < m^{k+1}$. By the division algorithm, we can divide n by m^k and obtain

$$n = a_k m^k + r, \quad \text{where } 0 \leq r < m^k.$$

Then

$$0 < m^k - r \leq n - r = a_k m^k \leq n < m^{k+1}.$$

Dividing this inequality by m^k, we obtain $0 < a_k < m$. Since m and a_k are integers, it follows that

$$1 \leq a_k \leq m - 1.$$

If $r = 0$, then $n = a_k m^k$ is an m-adic representation. If $r \geq 1$, then $m^{k'} \leq r < m^{k'+1}$ for some nonnegative integer $k' \leq k-1$. By the induction assumption, $S(k')$ is true and r has a unique m-adic representation of the form

$$r = a_0 + a_1 m + \cdots + a_{k-1} m^{k-1}$$

with $0 \leq a_i \leq m - 1$ for $i = 0, 1, \ldots, k-1$. It follows that n has the m-adic representation

$$n = a_0 + a_1 m + \cdots + a_{k-1} m^{k-1} + a_k m^k.$$

We shall show that this representation is unique. Let

$$n = b_0 + b_1 m + \cdots + b_\ell m^\ell$$

be another m-adic representation of n, where $0 \leq b_j \leq m - 1$ for all $j = 0, 1, \ldots, \ell$ and $b_\ell \geq 1$. If $\ell \geq k + 1$, then

$$n < m^{k+1} \leq b_\ell m^\ell \leq n,$$

which is impossible. If $\ell \leq k - 1$, then the inequalities $b_j \leq m - 1$ imply that

$$\begin{aligned} n &= b_0 + b_1 m + \cdots + b_\ell m^\ell \\ &\leq (m-1) + (m-1)m + \cdots + (m-1)m^\ell \\ &= m^{\ell+1} - 1 \\ &< m^k \\ &\leq n, \end{aligned}$$

which is also impossible. Therefore, $k = \ell$. If $a_k < b_k$, then

$$\begin{aligned} n &= a_0 + a_1 m + \cdots + a_{k-1} m^{k-1} + a_k m^k \\ &\leq (m-1) + (m-1)m + \cdots + (m-1)m^{k-1} + a_k m^k \\ &= (m^k - 1) + a_k m^k \\ &< (a_k + 1)m^k \\ &\leq b_k m^k \\ &\leq n, \end{aligned}$$

which again is impossible. Therefore, $b_k \leq a_k$. By symmetry, we have $a_k \leq b_k$ and so $a_k = b_k$. Then

$$\begin{aligned} n - a_k m^k &= a_0 + a_1 m + a_2 m^2 + \cdots + a_{k-1} m^{k-1} \\ &= b_0 + b_1 m + b_2 m^2 + \cdots + b_{k-1} m^{k-1} \\ &< m^k. \end{aligned}$$

By the induction assumption, $a_i = b_i$ for $i = 0, 1, \ldots, k-1$. Thus, the m-adic representation of n exists and is unique, and $S(k)$ is true. By mathematical induction, $S(k)$ holds for all $k \geq 0$. □

For example, the 2-adic representation of 100 is

$$100 = 1 \cdot 2^2 + 1 \cdot 2^5 + 1 \cdot 2^6,$$

and the 3-adic representation of 100 is

$$100 = 1 + 2 \cdot 3^2 + 1 \cdot 3^4.$$

The 10-adic representation of 217 is

$$217 = 7 + 1 \cdot 10^1 + 2 \cdot 10^2.$$

Exercises

1. Find all divisors of 20.

2. Find all divisors of 29,601.

3. Find all divisors of 1.

4. Find the quotient and remainder for a divided by d when

    (a) $a = 281$ and $d = 23$.
    (b) $a = 281$ and $d = 12$.
    (c) $a = 291$ and $d = 23$.
    (d) $a = 291$ and $d = 12$.

5. Find the quotient and remainder for $10^k + 1$ divided by 11 for $k = 1, 2, 3, 4, 5$.

6. Compute the m-adic representation of 526 for $m = 2, 3, 7$, and 9.

7. Compute the 100-adic representation of 783,614,955.

8. Prove that n is even, then n^2 is divisible by 4.

9. Prove that n is odd, then $n^2 - 1$ is divisible by 8.

10. Prove that $n^3 - n$ is divisible by 6 for every integer n.

11. Prove that if d divides a, then d^k divides a^k for every positive integer k.

12. Prove that if d divides a and d divides b, then d divides $ax + by$ for all integers x and y.

13. Prove that if a and d are integers such that d divides a and $|a| < d$, then $a = 0$.

14. Prove that divisibility is transitive, that is, if a divides b and b divides c, then a divides c.

15. Prove by induction that $n \leq 2^{n-1}$ for all positive integers n.

16. Prove by induction that

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

for all positive integers n.

17. Prove by induction that

$$1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$$

for all positive integers n, that is, the sum of the cubes of the first n integers is equal to the square of the sum of the first n integers.

18. Prove that the principle of mathematical induction is equivalent to the minimum principle.

19. Let a and d be integers with $d \geq 1$. Prove that there exist unique integers q' and r' such that

$$a = dq' + r'$$

and

$$-\frac{d}{2} < r' \leq \frac{d}{2}.$$

20. For integers n and k with $n \geq 1$ and $0 \leq k \leq n$, we define the *binomial coefficient*

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Define $\binom{0}{0} = 1$. Prove that for all $n \geq 1$,

$$\binom{n}{0} = \binom{n}{n} = 1$$

and

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

for $1 \le k \le n-1$.

21. Prove that the product of any k consecutive integers is always divisible by $k!$.

    Hint: Use induction on n to show that $\binom{n}{k}$ is an integer.

22. Let $m_0, m_1, m_2, \ldots$ be a strictly increasing sequence of positive integers such that $m_0 = 1$ and m_i divides m_{i+1} for all $i \ge 0$. Prove that every positive integer n can be represented uniquely in the form

$$n = \sum_{i=0}^{\infty} a_i m_i,$$

where

$$0 \le a_i \le \frac{m_{i+1}}{m_i} - 1 \qquad \text{for all } i \ge 0$$

and $m_i = 0$ for all but finitely many integers i.

23. Prove that every positive integer n can be represented uniquely in the form

$$n = \sum_{k=0}^{\infty} a_k k!,$$

where

$$0 \le a_k \le k.$$

24. Prove that every positive integer n can be uniquely represented in the form

$$n = b_0 + b_1 3 + b_2 3^2 + \cdots + b_{k-1} 3^{k-1} + 3^k,$$

where $b_i \in \{0, 1, -1\}$ for $i = 0, 1, 2, \ldots, k-1$.

25. Let $\mathbf{N}^k$ denote the set of all k-tuples of positive integers. We define the *lexicographic order* on $\mathbf{N}^k$ as follows. For $(a_1, \ldots, a_k), (b_1, \ldots, b_k) \in \mathbf{N}^k$, we write

$$(a_1, \ldots, a_k) \preceq (b_1, \ldots, b_k)$$

if either $a_i = b_i$ for all $i = 1, \ldots, k$, or there exists an integer j such that $a_i = b_i$ for $i < j$ and $a_j < b_j$. Prove that

    (a) The relation $\preceq$ is *reflexive* in the sense that if $(a_1, \ldots, a_k) \preceq (b_1, \ldots, b_k)$ and $(b_1, \ldots, b_k) \preceq (a_1, \ldots, a_k)$, then $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$.

(b) The relation $\preceq$ is *transitive* in the sense that if $(a_1, \ldots, a_k) \preceq (b_1, \ldots, b_k)$ and $(b_1, \ldots, b_k) \preceq (c_1, \ldots, c_k)$, then $(a_1, \ldots, a_k) \preceq (c_1, \ldots, c_k)$.

(c) The relation $\preceq$ is *total* in the sense that if $(a_1, \ldots, a_k), (b_1, \ldots, b_k) \in \mathbf{N}^k$, then $(a_1, \ldots, a_k) \preceq (b_1, \ldots, b_k)$ or $(b_1, \ldots, b_k) \preceq (a_1, \ldots, a_k)$.

A relation that is reflexive and transitive is called a *partial order.* A partial order that is total is called a *total order.* Thus, the lexicographic order is a total order on the set of k-tuples of positive integers.

26. Prove that $\mathbf{N}^k$ with the lexicographic order satisfies the following minimum principle: Every nonempty set of k-tuples of positive integers contains a smallest element.

1.2 Greatest Common Divisors

Algebra is a natural language to describe many results in elementary number theory.

Let G be a nonempty set, and let $G \times G$ denote the set of all ordered pairs (x, y) with $x, y \in G$. A *binary operation* on G is a map from $G \times G$ into G. We denote the image of $(x, y) \in G \times G$ by $x * y \in G$.

A *group* is a set G with a binary operation that satisfies the following three axioms:

(i) Associativity: For all $x, y, z \in G$,

$$(x * y) * z = x * (y * z).$$

(ii) Identity element: There exists an element $e \in G$ such that for all $x \in G$,

$$e * x = x * e = x.$$

The element e is called the *identity* of the group.

(iii) Inverses: For every $x \in G$ there exists an element $y \in G$ such that

$$x * y = y * x = e.$$

The element y is called the *inverse* of x.

The group G is called *abelian* or *commutative* if the binary operation also satisfies the axiom

(iv) Commutativity: For all $x, y \in G$,

$$x * y = y * x.$$

We can use additive notation and denote the image of the ordered pair $(x, y) \in G \times G$ by $x + y$. We call $x + y$ the *sum* of x and y. In an additive group, the identity is usually written 0, the inverse of x is written $-x$, and we define $x - y = x + (-y)$. We can also use multiplicative notation and denote the image of the ordered pair $(x, y) \in G \times G$ by xy. We call xy the *product* of x and y. In a multiplicative group, the identity is usually written 1 and the inverse of x is written x^{-1}.

Examples of abelian groups are the integers $\mathbf{Z}$, the rational numbers $\mathbf{Q}$, the real numbers $\mathbf{R}$, and the complex numbers $\mathbf{C}$, with the usual operation of addition. The nonzero rational, real, and complex numbers, denoted by $\mathbf{Q}^\times, \mathbf{R}^\times$, and $\mathbf{C}^\times$, respectively, are also abelian groups, with the usual multiplication as the binary operation. For every positive integer m, the set of complex numbers

$$\Gamma_m = \{e^{2\pi ik/m} : k = 0, 1, \ldots, m - 1\}$$

is a multiplicative group. The elements of Γ_m are called *mth roots of unity*, since $\omega^m = 1$ for all $\omega \in \Gamma_m$. An example of a nonabelian group is the set $GL_2(\mathbf{C})$ of 2×2 matrices with complex coefficients and nonzero determinant, and with the usual matrix multiplication as the binary operation.

A *subgroup* of a group G is a nonempty subset of G that is also a group under the same binary operation as G. If H is a subgroup of G, then H is closed under the binary operation in G, H contains the identity element of G, and the inverse of every element of H belongs to H. For example, the set of even integers is a subgroup of $\mathbf{Z}$. A nonempty subset H of an additive abelian group G is a subgroup if and only if $x - y \in H$ for all $x, y \in H$ (Exercise 20).

For every integer d, the set of all multiples of d is a subgroup of $\mathbf{Z}$. We denote this subgroup by $d\mathbf{Z}$. If $a_1, \ldots, a_k \in \mathbf{Z}$, then the set of all numbers of the form $a_1x_1 + \cdots + a_kx_k$ with $x_1, \ldots, x_k \in \mathbf{Z}$ is also a subgroup of $\mathbf{Z}$. The set $\mathbf{Q}$ of rational numbers is a subgroup of the additive group $\mathbf{R}$. The set $\mathbf{R}^+$ of positive real numbers is a subgroup of the multiplicative group $\mathbf{R}^\times$. Let $\mathbf{T} = \{z \in \mathbf{C} : |z| = 1\}$ denote the set of complex numbers of absolute value 1, that is, the unit circle in the complex plane. Then $\mathbf{T}$ is a subgroup of the multiplicative group $\mathbf{C}^\times$, and Γ_m is a subgroup of $\mathbf{T}$.

If G is a group, written multiplicatively, and $g \in G$, then $g^n \in G$ for all $n \in \mathbf{Z}$ (Exercise 21), and $\{g^n : n \in \mathbf{Z}\}$ is a subgroup of G.

The intersection of a family of subgroups of a group G is a subgroup of G (Exercise 22). Let S be a subset of a group G. The *subgroup of G generated by S* is the smallest subgroup of G that contains S. This is simply the intersection of all subgroups of G that contain S (Exercise 23). For example, the subgroup of $\mathbf{Z}$ generated by the set $\{d\}$ is $d\mathbf{Z}$.

Theorem 1.3 *Let H be a subgroup of the integers under addition. There exists a unique nonnegative integer d such that H is the set of all multiples*

of d, that is,

$$H = \{0, \pm d, \pm 2d, \ldots\} = d\mathbf{Z}.$$

Proof. We have $0 \in H$ for every subgroup H. If $H = \{0\}$ is the zero subgroup, then we choose $d = 0$ and $H = 0\mathbf{Z}$. Moreover, $d = 0$ is the unique generator of this subgroup.

If $H \neq \{0\}$, then there exists $a \in H, a \neq 0$. Since $-a$ also belongs to H, it follows that H contains positive integers. By the minimum principle, H contains a least positive integer d. By Exercise 21, $dq \in H$ for every integer q, and so $d\mathbf{Z} \subseteq H$.

Let $a \in H$. By the division algorithm, we can write $a = dq + r$, where q and r are integers and $0 \leq r \leq d - 1$. Since $dq \in H$ and H is closed under subtraction, it follows that

$$r = a - dq \in H.$$

Since $0 \leq r < d$ and d is the smallest positive integer in H, we must have $r = 0$, that is, $a = dq \in d\mathbf{Z}$ and $H \subseteq d\mathbf{Z}$. It follows that $H = d\mathbf{Z}$.

If $H = d\mathbf{Z} = d'\mathbf{Z}$, where d and d' are positive integers, then $d' \in d\mathbf{Z}$ implies that $d' = dq$ for some integer q, and $d \in d'\mathbf{Z}$ implies that $d = d'q'$ for some integer q'. Therefore,

$$d = d'q' = dqq',$$

and so $qq' = 1$, hence $q = q' = \pm 1$ and $d = \pm d'$. Since d and d' are positive, we have $d = d'$, and d is the unique positive integer that generates the subgroup H. □

For example, if H is the subgroup consisting of all integers of the form $35x + 91y$, then $7 = 35(-5) + 91(2) \in H$ and $H = 7\mathbf{Z}$.

Let A be a nonempty set of integers, not all 0. If the integer d divides a for all $a \in A$, then d is called a *common divisor* of A. For example, 1 is a common divisor of every nonempty set of integers. The positive integer d is called the *greatest common divisor* of the set A, denoted by $d = \gcd(A)$, if d is a common divisor of A and every common divisor of A divides d. We shall prove that every nonempty set of integers has a greatest common divisor.

Theorem 1.4 *Let A be a nonempty set of integers, not all zero. Then A has a unique greatest common divisor, and there exist integers $a_1, \ldots, a_k \in A$ and $x_1, \ldots, x_k$ such that*

$$\gcd(A) = a_1x_1 + \cdots + a_kx_k.$$

Proof. Let H be the subset of $\mathbf{Z}$ consisting of all integers of the form

$$a_1x_1 + \cdots + a_kx_k \qquad \text{with } a_1, \ldots, a_k \in A \text{ and } x_1, \ldots, x_k \in \mathbf{Z}.$$

Then H is a subgroup of $\mathbf{Z}$ and $A \subseteq H$. By Theorem 1.3, there exists a unique positive integer d such that $H = d\mathbf{Z}$, that is, H consists of all multiples of d. In particular, every integer $a \in A$ is a multiple of d, and so d is a common divisor of A. Since $d \in H$, there exist integers $a_1, \ldots, a_k \in A$ and $x_1, \ldots, x_k$ such that

$$d = a_1x_1 + \cdots + a_kx_k.$$

It follows that every common divisor of A must divide d, hence d is a greatest common divisor of A.

If the positive integers d and d' are both greatest common divisors, then $d|d'$ and $d'|d$, and so $d = d'$. It follows that $\gcd(A)$ is unique. □

If $A = \{a_1, \ldots, a_k\}$ is a nonempty, finite set of integers, not all 0, we write $\gcd(A) = (a_1, \ldots, a_k)$. For example,

$$(35, 91) = 7 = 35(-5) + 91(2).$$

Theorem 1.5 *Let $a_1, \ldots, a_k$ be integers, not all zero. Then $(a_1, \ldots, a_k) = 1$ if and only if there exist integers $x_1, \ldots, x_k$ such that*

$$a_1x_1 + \cdots + a_kx_k = 1.$$

Proof. This follows immediately from Theorem 1.4. □

The integers $a_1, \ldots, a_k$ are called *relatively prime* if their greatest common divisor is 1, that is, $(a_1, \ldots, a_k) = 1$. The integers $a_1, \ldots, a_k$ are called *pairwise relatively prime* if $(a_i, a_j) = 1$ for $i \neq j$. For example, the three integers $6, 10, 15$ are relatively prime but not pairwise relatively prime, since $(6, 10, 15) = 1$ but $(6, 10) = 2$, $(6, 15) = 3$, and $(10, 15) = 5$.

Let G and H be groups, and denote the group operations by $*$. A map $f : G \rightarrow H$ is called a *group homomorphism* if $f(x * y) = f(x) * f(y)$ for all $x, y \in G$. Thus, a homomorphism f from an additive group G into a multiplicative group H is a map such that $f(x + y) = f(x)f(y)$ for all $x, y \in G$. For example, if $\mathbf{R}$ is the additive group of real numbers and $\mathbf{R}^+$ is the multiplicative group of positive real numbers, then the exponential map $\exp : \mathbf{R} \rightarrow \mathbf{R}^+$ defined by $\exp(x) = e^x$ is a homomorphism.

A group homomorphism $f : G \rightarrow H$ is called an *isomorphism* if f is one-to-one and onto. Groups G and H are called *isomorphic*, denoted by $G \cong H$, if there exists an isomorphism between them. For example, let $2\mathbf{Z}$ denote the additive group of even integers. The map $f : \mathbf{Z} \rightarrow 2\mathbf{Z}$ defined by $f(n) = 2n$ is an isomorphism between the group of integers and the subgroup of even integers.

Exercises

1. Compute $(935, 1122)$.

2. Compute $(168, 252, 294)$.

3. Find integers x and y such that $13x + 15y = 1$.

4. Construct four relatively prime integers a, b, c, d such that no three of them are relatively prime.

5. Prove that $(n, n+2) = 1$ is n is odd and $(n, n+2) = 2$ is n is even.

6. Prove that $2n+5$ and $3n+7$ are relatively prime for every integer n.

7. Prove that $3n+2$ and $5n+3$ are relatively prime for every integer n.

8. Prove that $n!+1$ and $(n+1)!+1$ are relatively prime for every positive integer n.

9. Let a, b, and d be positive integers. Prove that if $(a, b) = 1$ and d divides a, then $(d, b) = 1$.

10. Let a and b be positive integers. Prove that $(a, b) = a$ if and only if a divides b.

11. Let a, b, c be positive integers. Prove that
$$(ac, bc) = (a, b)c.$$

12. Let a, b, and c be positive integers. Prove that
$$((a, b), c) = (a, (b, c)) = (a, b, c).$$

13. Let A be a nonempty set of integers. Prove that the greatest common divisor of A is the largest integer that divides every element of A.

14. Let a, b, c, d be integers such that $ad - bc = 1$. For integers u and v, define
$$u' = au + bv$$
and
$$v' = cu + dv.$$
Prove that $(u, v) = (u', v')$.

Hint: Express u and v in terms of u' and v'.

15. Let $S = \mathbf{Q}^{n+1} \setminus \{(0,0,\ldots,0)\}$ denote the set of all nonzero $(n+1)$-tuples of rational numbers. If t is a nonzero rational number and $(x_0, x_1, \ldots, x_n) \in S$, then we define

$$t(x_0, x_1, \ldots, x_n) = (tx_0, tx_1, \ldots, tx_n) \in S.$$

We introduce a relation $\sim$ on S as follows: If $(x_0, x_1, \ldots, x_n)$ and $(y_0, y_1, \ldots, y_n)$ are in S, then $(x_0, x_1, \ldots, x_n) \sim (y_0, y_1, \ldots, y_n)$ if there exists a nonzero rational number t such that $t(x_0, x_1, \ldots, x_n) = (y_0, y_1, \ldots, y_n)$. Prove that this is an equivalence relation, that is, prove that $\sim$ is reflexive ($x \sim x$ for all $x \in S$), symmetric (if $x \sim y$, then $y \sim x$), and transitive (if $x \sim y$ and $y \sim z$, then $x \sim z$). The set of equivalence classes of this relation is called *n-dimensional projective space over the field of rational numbers*, and denoted by $\mathbf{P}^n(\mathbf{Q})$.

16. Consider $\left(\frac{25}{6}, -5, \frac{10}{3}\right) \in \mathbf{Q}^3$. Find all triples (a_0, a_1, a_2) of relatively prime integers such that

$$(a_0, a_1, a_2) \sim \left(\frac{25}{6}, -5, \frac{10}{3}\right).$$

17. Let

$$(x_0, x_1, \ldots, x_n) \in S = \mathbf{Q}^{n+1} \setminus \{(0,0,\ldots,0)\}.$$

Let $[(x_0, x_1, \ldots, x_n)]$ denote the equivalence class of $(x_0, x_1, \ldots, x_n)$ in $\mathbf{P}^n(\mathbf{Q})$. Prove that there exist exactly two elements $(a_0, a_1, \ldots, a_n)$ and $(b_0, b_1, \ldots, b_n)$ in S such that the numbers $a_0, a_1, \ldots, a_n$ are relatively prime integers, the numbers $b_0, b_1, \ldots, b_n$ are relatively prime integers, and

$$[(x_0, x_1, \ldots, x_n)] = [(a_0, a_1, \ldots, a_n)] = [(b_0, b_1, \ldots, b_n)] \in \mathbf{P}^n(\mathbf{Q}).$$

Moreover,

$$(b_0, b_1, \ldots, b_n) = -(a_0, a_1, \ldots, a_n).$$

18. Prove that the set of all rational numbers of the form $a/2^k$, where $a \in \mathbf{Z}$ and $k \in \mathbf{N}_0$, is an additive subgroup of $\mathbf{Q}$.

19. Let $G = \{2\mathbf{Z}, 1+2\mathbf{Z}\}$, where $2\mathbf{Z}$ denotes the set of even integers and $1+2\mathbf{Z}$ the set of odd integers. Define addition of elements of G by

$$2\mathbf{Z} + 2\mathbf{Z} = (1+2\mathbf{Z}) + (1+2\mathbf{Z}) = 2\mathbf{Z}$$

and

$$2\mathbf{Z} + (1+2\mathbf{Z}) = (1+2\mathbf{Z}) + 2\mathbf{Z} = 1+2\mathbf{Z}.$$

Prove that G is an additive abelian group.

20. Let H be a nonempty subset of an additive abelian group G. Prove that H is a subgroup if and only if $x - y \in H$ for all $x, y \in H$.

21. Prove that if G is a group, written multiplicatively, and $g \in G$, then $g^n \in G$ for all $n \in \mathbf{Z}$. (If G is an additive group, then $ng \in G$ for all $n \in \mathbf{Z}$.)

22. Prove that the intersection of a family of subgroups of a group G is a subgroup of G.

23. Let S be a nonempty subset of an additive abelian group G. Prove that the subgroup of G generated by S is the intersection of all subgroups of G that contain S.

24. Prove that every nonzero subgroup of $\mathbf{Z}$ is isomorphic to $\mathbf{Z}$.

25. Let G be the set of all matrices of the form

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix},$$

with $a \in \mathbf{Z}$ and matrix multiplication as the binary operation. Prove that G is an abelian group isomorphic to $\mathbf{Z}$.

26. Let $H_3(\mathbf{Z})$ be the set of all matrices of the form

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

with $a, b, c \in \mathbf{Z}$ and matrix multiplication as the binary operation. Prove that $H_3(\mathbf{Z})$ is a nonabelian group. This group is called the *Heisenberg group.*

27. Let $\mathbf{R}$ be the additive group of real numbers and $\mathbf{R}^+$ the multiplicative group of positive real numbers. Let $\exp : \mathbf{R} \to \mathbf{R}^+$ be the exponential map $\exp(x) = e^x$. Prove that the exponential map is a group isomorphism.

28. Let G and H be groups with e the identity in H. Let $f : G \to H$ be a group homomorphism. The *kernel* of f is the set

$$f^{-1}(e) = \{x \in G : f(x) = e \in H\} \subseteq G.$$

The *image* of f is the set

$$f(G) = \{f(x) : x \in G\} \subseteq H.$$

Prove that the kernel of f is a subgroup of G, and the image of f is a subgroup of H.

29. Define the map $f : \mathbf{Z} \to \mathbf{Z}$ by $f(n) = 3n$. Prove that f is a group homomorphism and determine the kernel and image of f.

30. Let Γ_m denote the multiplicative group of mth roots of unity. Prove that the map $f : \mathbf{Z} \to \Gamma_m$ defined by $f(k) = e^{2\pi ik/m}$ is a group homomorphism. What is the kernel of this homomorphism?

31. Let $G = [0, 1)$ be the interval of real numbers x such that $0 \leq x < 1$. We define a binary operation $x * y$ for numbers $x, y \in G$ as follows:

$$x * y = \begin{cases} x + y & \text{if } x + y < 1, \\ x + y - 1 & \text{if } x + y \geq 1. \end{cases}$$

Prove that G is an abelian group with this operation. This group is denoted by $\mathbf{R}/\mathbf{Z}$.

Define the map $f : \mathbf{R} \to \mathbf{R}/\mathbf{Z}$ by $f(t) = \{t\}$, where $\{t\}$ denotes the fractional part of t. Prove that f is a group homomorphism. What is the kernel of this homomorphism?

1.3 The Euclidean Algorithm and Continued Fractions

Let a and b be integers with $b \geq 1$. There is a simple and efficient method to compute the greatest common divisor of a and b and to express (a, b) explicitly in the form $ax + by$. Define $r_0 = a$ and $r_1 = b$. By the division algorithm, there exist integers q_0 and r_2 such that

$$r_0 = r_1 q_0 + r_2$$

and

$$0 \leq r_2 < r_1.$$

If an integer d divides r_0 and r_1, then d also divides r_1 and r_2. Similarly, if an integer d divides r_1 and r_2, then d also divides r_0 and r_1. Therefore, the set of common divisors of r_0 and r_1 is the same as the set of common divisors of r_1 and r_2, and so

$$(a, b) = (r_0, r_1) = (r_1, r_2).$$

If $r_2 = 0$, then $a = bq_0$ and $(a, b) = b = r_1$. If $r_2 > 0$, then we divide r_2 into r_1 and obtain integers q_1 and r_3 such that

$$r_1 = r_2 q_1 + r_3,$$

where

$$0 \leq r_3 < r_2 < r_1$$

and

$$(a, b) = (r_1, r_2) = (r_2, r_3).$$

Moreover, $q_1 \geq 1$ since $r_2 < r_1$. If $r_3 = 0$, then $(a, b) = r_2$. If $r_3 > 0$, then there exist integers q_2 and r_4 such that

$$r_2 = r_3 q_2 + r_4,$$

where $q_2 \geq 1$ and

$$0 \leq r_4 < r_3 < r_2 < r_1$$

and

$$(a, b) = (r_2, r_3) = (r_3, r_4).$$

If $r_4 = 0$, then $(a, b) = r_3$.

Iterating this process k times, we obtain an integer q_0, a sequence of positive integers $q_1, q_2, \ldots, q_{k-1}$, and a strictly decreasing sequence of nonnegative integers $r_1, r_2, \ldots, r_{k+1}$ such that

$$r_{i-1} = r_i q_{i-1} + r_{i+1}$$

for $i = 1, 2, \ldots, k$, and

$$(a, b) = (r_0, r_1) = (r_1, r_2) = \cdots = (r_k, r_{k+1}).$$

If $r_{k+1} > 0$, then we can divide r_k by r_{k+1} and obtain

$$r_k = r_{k+1} q_k + r_{k+2},$$

where $0 \leq r_{k+2} < r_{k+1}$. Since a strictly decreasing sequence of nonnegative integers must be finite, it follows that there exists an integer $n \geq 1$ such that $r_{n+1} = 0$. Then we have an integer q_0, a sequence of positive integers $q_1, q_2, \ldots, q_{n-1}$, and a strictly decreasing sequence of positive integers $r_1, r_2, \ldots, r_n$ with

$$(a, b) = (r_n, r_{n+1}) = r_n.$$

The n applications of the division algorithm produce n equations

$$\begin{aligned}
r_0 &= r_1 q_0 + r_2 \\
r_1 &= r_2 q_1 + r_3 \\
r_2 &= r_3 q_2 + r_4 \\
&\vdots \\
r_{n-2} &= r_{n-1} q_{n-2} + r_n \\
r_{n-1} &= r_n q_{n-1}.
\end{aligned}$$

Since $r_n < r_{n+1}$, it follows that $q_{n-1} \geq 2$.

This procedure is called the *Euclidean algorithm*. We call n the *length* of the Euclidean algorithm for a and b. This is the number of divisions

required to find the greatest common divisor. The sequence $q_0, q_1, \ldots, q_{n-1}$ is called the *sequence of partial quotients*. The sequence $r_2, r_3, \ldots, r_n$ is called the *sequence of remainders.*

Let us use the Euclidean algorithm to find $(574, 252)$ and express it as a linear combination of 574 and 252. We have

$$\begin{aligned} 574 &= 252 \cdot 2 + 70, \\ 252 &= 70 \cdot 3 + 42, \\ 70 &= 42 \cdot 1 + 28, \\ 42 &= 28 \cdot 1 + 14, \\ 28 &= 14 \cdot 2, \end{aligned}$$

and so

$$(574, 252) = 14.$$

The sequence of partial quotients is $(2, 3, 1, 1, 2)$ and the sequence of partial remainders is $(70, 42, 28, 14)$. The Euclidean algorithm for 574 and 252 has length 5. Note that $574 = 14 \cdot 41$ and $252 = 14 \cdot 18$, and that 41 and 18 are relatively prime. Working backwards through the Euclidean algorithm to express 14 as a linear combination of 574 and 252, we obtain

$$\begin{aligned} 14 &= 42 - 28 \cdot 1 \\ &= 42 - (70 - 42 \cdot 1) \cdot 1 = 42 \cdot 2 - 70 \cdot 1 \\ &= (252 - 70 \cdot 3) \cdot 2 - 70 \cdot 1 = 252 \cdot 2 - 70 \cdot 7 \\ &= 252 \cdot 2 - (574 - 252 \cdot 2) \cdot 7 = 252 \cdot 16 - 574 \cdot 7. \end{aligned}$$

Let $a_0, a_1, \ldots, a_N$ be real numbers with $a_i > 0$ for $i = 1, \ldots, N$. We define the *finite simple continued fraction*

$$\langle a_0, a_1, \ldots, a_N \rangle = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \cfrac{1}{a_{N-1} + \cfrac{1}{a_N}}}}}.$$

Another notation for a continued fraction is

$$\langle a_0, a_1, \ldots, a_N \rangle = a_0 + \frac{1}{a_1 +} \frac{1}{a_2 +} \cdots \frac{1}{a_N}.$$

The numbers $a_0, a_1, \ldots, a_N$ are called the *partial quotients* of the continued fraction. For example,

$$\langle 2, 1, 1, 2 \rangle = 2 + \cfrac{1}{1 + \cfrac{1}{1 + \frac{1}{2}}} = \frac{13}{5}.$$

We can write a finite simple continued fraction as a rational function in the variables $a_0, a_1, \ldots, a_N$. For example,

$$\langle a_0 \rangle = a_0,$$

$$\langle a_0, a_1 \rangle = \frac{a_0 a_1 + 1}{a_1},$$

and

$$\langle a_0, a_1, a_2 \rangle = \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1}.$$

If $N \geq 1$, then (Exercise 5)

$$\langle a_0, a_1, \ldots, a_N \rangle = a_0 + \frac{1}{\langle a_1, \ldots, a_N \rangle}.$$

We can use the Euclidean algorithm to write a rational number as a finite simple continued fraction with integral partial quotients. For example, to represent 574/274, we have

$$\begin{aligned}
\frac{574}{252} &= 2 + \frac{70}{252} \\
&= 2 + \frac{1}{3 + \frac{42}{70}} \\
&= 2 + \frac{1}{3 + \frac{1}{1 + \frac{28}{42}}} \\
&= 2 + \frac{1}{3 + \frac{1}{1 + \frac{1}{1 + \frac{14}{28}}}} \\
&= 2 + \frac{1}{3 + \frac{1}{1 + \frac{1}{1 + \frac{1}{2}}}} \\
&= \langle 2, 3, 1, 1, 2 \rangle.
\end{aligned}$$

Notice that the partial quotients in the Euclidean algorithm are the partial quotients in the continued fraction.

Theorem 1.6 *Let a and b be integers with $b \geq 1$. If the Euclidean algorithm for a and b has length n with sequence of partial quotients $q_0, q_1, \ldots, q_{n-1}$, then*

$$\frac{a}{b} = \langle q_0, q_1, \ldots, q_{n-1} \rangle.$$

Proof. Let $r_0 = a$ and $r_1 = b$. The proof is by induction on n. If $n = 1$, then

$$r_0 = r_1 q_0$$

and

$$\frac{a}{b} = \frac{r_0}{r_1} = q_0 = \langle q_0 \rangle.$$

If $n = 2$, then

$$\begin{aligned}
r_0 &= r_1 q_0 + r_2, \\
r_1 &= r_2 q_1,
\end{aligned}$$

and

$$\frac{a}{b} = \frac{r_0}{r_1} = q_0 + \frac{r_2}{r_1} = q_0 + \frac{1}{\frac{r_1}{r_2}} = q_0 + \frac{1}{q_1} = \langle q_0, q_1 \rangle.$$

Let $n \geq 2$, and assume that the theorem is true for integers a and $b \geq 1$ whose Euclidean algorithm has length n. Let a and $b \geq 1$ be integers whose Euclidean algorithm has length $n+1$ and whose sequence of partial quotients is $\langle q_0, q_1, \ldots, q_n \rangle$. Let

$$\begin{aligned} r_0 &= r_1 q_0 + r_2 \\ r_1 &= r_2 q_1 + r_3 \\ &\vdots \\ r_{n-1} &= r_n q_{n-1} + r_{n+1} \\ r_n &= r_{n+1} q_n. \end{aligned}$$

be the $n+1$ equations in the Euclidean algorithm for $a = r_0$ and $b = r_1$. The Euclidean algorithm for the positive integers r_1 and r_2 has length n with sequence of partial quotients $q_1, \ldots, q_n$. It follows from the induction hypothesis that

$$\frac{r_1}{r_2} = \langle q_1, \ldots, q_n \rangle$$

and so

$$\frac{a}{b} = \frac{r_0}{r_1} = q_0 + \frac{1}{\frac{r_1}{r_2}} = q_0 + \frac{1}{\langle q_1, \ldots, q_n \rangle} = \langle q_0, q_1, \ldots, q_n \rangle.$$

This completes the proof. □

It is also true that the representation of a rational number as a finite simple continued fraction is essentially unique (Exercise 8).

Exercises

1. Use the Euclidean algorithm to compute the greatest common divisor of 35 and 91, and to express $(35, 91)$ as a linear combination of 35 and 91. Compute the simple continued fraction for 91/35.

2. Use the Euclidean algorithm to write the greatest common divisor of 4534 and 1876 as a linear combination of 4534 and 1876. Compute the simple continued fraction for 4534/1876.

3. Use the Euclidean algorithm to compute the greatest common divisor of 1197 and 14280, and to express $(1197, 14280)$ as a linear combination of 1197 and 14280.

4. Compute the simple continued fraction $\langle 2,1,2,1,1,4\rangle$ to 4 decimal places, and compare this number to e.

5. Prove that
$$\langle a_0, a_1, \ldots, a_N\rangle = a_0 + \frac{1}{\langle a_1, \ldots, a_N\rangle}.$$

6. Let $N \geq 1$. Prove that
$$\langle a_0, a_1, \ldots, a_{N-2}, a_{N-1}, 1\rangle = \langle a_0, a_1, \ldots, a_{N-2}, a_{N-1}+1\rangle.$$

7. Let $x = \langle a_0, a_1, \ldots, a_N\rangle$ be a finite simple continued fraction whose partial quotients a_i are integers, with $N \geq 1$ and $a_N \geq 2$. Let $[x]$ denote the integer part of x and $\{x\}$ the fractional part of x. Prove that
$$[x] = a_0$$
and
$$\{x\} = \frac{1}{\langle a_1, \ldots, a_N\rangle}.$$

8. Let $\frac{a}{b}$ be a rational number that is not an integer. Prove that there exist unique integers $a_0, a_1, \ldots, a_N$ such that $a_i \geq 1$ for $i = 1, \ldots, N-1$, $a_N \geq 2$, and
$$\frac{a}{b} = \langle a_0, a_1, \ldots, a_{N-1}, a_N\rangle.$$
Hint: By Exercise 7, if
$$x = \langle a_0, a_1, \ldots, a_N\rangle = \langle b_0, b_1, \ldots, b_M\rangle$$
with $a_i, b_j \in \mathbf{Z}$ and $a_N, b_M \geq 2$, then $a_0 = [x] = b_0$.

9. Prove that
$$\langle a_0, a_1, \ldots, a_N, a_{N+1}\rangle = \langle a_0, a_1, \ldots, a_N + \frac{1}{a_{N+1}}\rangle.$$

10. Let $\langle a_0, a_1, \ldots, a_N\rangle$ be a finite simple continued fraction. Define
$$p_0 = a_0,$$
$$p_1 = a_1 a_0 + 1,$$
and
$$p_n = a_n p_{n-1} + p_{n-2} \qquad \text{for } n = 2, \ldots, N.$$
Define
$$q_0 = 1,$$
$$q_1 = a_1,$$

and

$$q_n = a_n q_{n-1} + q_{n-2} \qquad \text{for } n = 2, \ldots, N.$$

Prove that

$$\langle a_0, a_1, \ldots, a_n \rangle = \frac{p_n}{q_n}$$

for $n = 0, 1, \ldots, N$. The continued fraction $\langle a_0, a_1, \ldots, a_n \rangle$ is called the nth *convergent* of the continued fraction $\langle a_0, a_1, \ldots, a_N \rangle$.

11. Compute the convergents p_n/q_n of the simple continued fraction $\langle 1, 2, 2, 2, 2, 2, 2 \rangle$. Compute p_6/q_6 to 5 decimal places, and compare this number to $\sqrt{2}$.

12. Let $\langle a_0, a_1, \ldots, a_N \rangle$ be a finite simple continued fraction, and let p_n and q_n be the numbers defined in Exercise 10. Prove that

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$$

and for $n = 1, \ldots, N$. Prove that if $a_i \in \mathbf{Z}$ for $i = 0, 1, \ldots, N$, then $(p_n, q_n) = 1$ for $n = 0, 1, \ldots, N$.

13. Let $\langle a_0, a_1, \ldots, a_N \rangle$ be a finite simple continued fraction, and let p_n and q_n be the numbers defined in Exercise 10. Prove that

$$p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n$$

for $n = 2, \ldots, N$.

14. Let $x = \langle a_0, a_1, \ldots, a_N \rangle$ be a finite simple continued fraction, and let p_n and q_n be the numbers defined in Exercise 10. Prove that the even convergents are strictly increasing, the odd convergents are strictly decreasing, and every even convergent is less than every odd convergent, that is,

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots \le x \le \cdots \frac{p_5}{q_5} < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

15. We define a sequence of integers as follows:

$$\begin{aligned} f_0 &= 0, \\ f_1 &= 1, \\ f_n &= f_{n-1} + f_{n-2} \qquad \text{for } n \ge 2. \end{aligned}$$

The integer f_n is called the nth *Fibonacci number*. Compute the Fibonacci numbers f_n for $n = 2, 3, \ldots, 12$. Prove that $(f_n, f_{n+1}) = 1$ for all nonnegative integers n.

In Exercises 16–23, f_n denotes the nth Fibonacci number.

16. Compute the convergents p_n/q_n of the simple continued fraction $\langle 1, 1, 1, 1, 1, 1, 1\rangle$. Observe that

$$\frac{p_n}{q_n} = \frac{f_{n+1}}{f_n}$$

for $n = 0, 1, \ldots, 6$.

17. Prove that

$$f_1 + f_2 + \cdots + f_n = f_{n+2} - 1$$

for all positive integers n.

18. Prove that

$$f_{n+1}f_{n-1} - f_n^2 = (-1)^n$$

for all positive integers n.

19. Prove that

$$f_n = f_{k+1}f_{n-k} + f_k f_{n-k-1}$$

for all $k = 0, 1, \ldots, n$. Equivalently,

$$\begin{aligned} f_n &= f_{n-1} + f_{n-2} = 2f_{n-2} + f_{n-3} \\ &= 3f_{n-3} + 2f_{n-4} = 5f_{n-4} + 3f_{n-5} \cdots . \end{aligned}$$

20. Prove that f_n divides $f_{\ell n}$ for all positive integers ℓ.

21. Prove that, for $n \geq 1$,

$$\begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n .$$

22. Let

$$\alpha = \frac{1 + \sqrt{5}}{2}$$

and

$$\beta = \frac{1 - \sqrt{5}}{2}.$$

Prove that

$$f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}} \qquad \text{for all } n \geq 0.$$

Prove that

$$f_n \sim \frac{\alpha^n}{\sqrt{5}} \qquad \text{as } n \to \infty$$

and

$$f_n \geq \alpha^{n-2} \quad \text{for } n \geq 2.$$

23. (Lamé's theorem) Let a and b be positive integers with $a > b$. The *length of the Euclidean algorithm* for a and b, denoted by $E(a,b)$, is the number of divisions required to find the greatest common divisor of a and b. Prove that

$$E(a,b) \leq \frac{\log b}{\log \alpha} + 1,$$

where $\alpha = (1+\sqrt{5})/2$.

Hint: Let $n = E(a,b)$. Set $r_0 = a$ and $r_1 = b$. For $i = 1, \ldots, n$, let

$$r_{i-1} = r_i q_{i-1} + r_{i+1},$$

where the positive integers $q_0, q_1, \ldots, q_{n-1}$ are the partial quotients and $r_2, \ldots, r_{n-1}, r_n$ are the remainders in the Euclidean algorithm. Then

$$r_0 > r_1 > \cdots > r_{n-1} > r_n \geq 1$$

and $(a,b) = (r_0, r_1) = r_n$. Let f_n be the nth Fibonacci number. Since $r_n \geq 1 = f_2$ and $r_{n-1} \geq 2 = f_3$, it follows that

$$\begin{aligned} r_{n-2} &= r_{n-1}q_{n-2} + r_n \geq f_3 + f_2 = f_4, \\ r_{n-3} &= r_{n-2}q_{n-3} + r_{n-1} \geq f_4 + f_3 = f_5, \end{aligned}$$

and, by induction on k,

$$r_{n-k} \geq f_{k+2}$$

for $k = 0, 1, \ldots, n$. In particular,

$$b = r_1 \geq f_{n+1} \geq \alpha^{n-1}.$$

1.4 The Fundamental Theorem of Arithmetic

A *prime number* is an integer p greater than 1 whose only positive divisors are 1 and p. A positive integer greater than 1 that is not prime is called *composite*. If n is composite, then it has a divisor d such that $1 < d < n$, and so $n = dd'$, where also $1 < d' < n$. The primes less than 100 are the following:

2	3	5	7	11
13	17	19	23	29
31	37	41	43	47
53	59	61	67	71
73	79	83	89	97.

If d is a positive divisor of n, then $d' = n/d$ is called the *conjugate divisor* to d. If $n = dd'$ and $d \leq d'$, then $d \leq \sqrt{n}$.

We shall prove that every positive integer can be written as the product of prime numbers (with the convention that the empty product is equal to 1), and that this representation is unique except for the order in which the prime factors are written. This result is called the *fundamental theorem of arithmetic.*

Theorem 1.7 (Euclid's lemma) *Let a, b, c be integers. If a divides bc and $(a, b) = 1$, then a divides c.*

Proof. Since a divides bc, we have $bc = aq$ for some integer q. Since a and b are relatively prime, Theorem 1.5 implies that there exist integers x and y such that

$$1 = ax + by.$$

Multiplying by c, we obtain

$$c = acx + bcy = acx + aqy = a(cx + qy),$$

and so a divides c. This completes the proof. □

Theorem 1.8 *Let $k \geq 2$, and let $a, b_1, b_2, \ldots, b_k$ be integers. If $(a, b_i) = 1$ for all $i = 1, \ldots, k$, then $(a, b_1 b_2 \cdots b_k) = 1$.*

Proof. The proof is by induction on k. Let $k = 2$ and $d = (a, b_1 b_2)$. We must show that $d = 1$. Since d divides a and $(a, b_1) = 1$, it follows that $(d, b_1) = 1$. Since d divides $b_1 b_2$, Euclid's lemma implies that d divides b_2. Therefore, d is a common divisor of a and b_2, but $(a, b_2) = 1$ and so $d = 1$.

Let $k \geq 3$, and assume that the result holds for $k - 1$. Let $a, b_1, \ldots, b_k$ be integers such that $(a, b_i) = 1$ for $i = 1, \ldots, k$. The induction assumption implies that $(a, b_1 \cdots b_{k-1}) = 1$. Since we also have $(a, b_k) = 1$, it follows from the case $k = 2$ that $(a, b_1 \cdots b_{k-1} b_k) = 1$. This completes the proof. □

Theorem 1.9 *If a prime number p divides a product of integers, then p divides one of the factors.*

Proof. Let $b_1, b_2, \ldots, b_k$ be integers such that p divides $b_1 \cdots b_k$. By Theorem 1.8, we have $(p, b_i) > 1$ for some i. Since p is prime, it follows that p divides b_i. □

Theorem 1.10 (Fundamental theorem of arithmetic) *Every positive integer can be written uniquely (up to order) as the product of prime numbers.*

Proof. First we prove that every positive integer can be written as a product of primes. Since an empty product is equal to 1, we can write 1 as the empty product of primes. Let $n \geq 2$. Suppose that every positive integer less than n is a product of primes. If n is prime, we are done. If n is composite, then $n = dd'$, where $1 < d \leq d' < n$. By the induction hypothesis, d and d' are both products of primes, and so $n = dd'$ is a product of primes.

Next we use induction to prove that this representation is unique. The representation of 1 as the product of the empty set of primes is unique. Let $n \geq 2$ and assume that the statement is true for all positive integers less than n. We must show that if $n = p_1 \cdots p_k = p'_1 \cdots p'_\ell$, where $p_1, \ldots, p_k, p'_1, \ldots, p'_\ell$ are primes, then $k = \ell$ and there is a permutation σ of $1, \ldots, k$ such that $p_i = p'_{\sigma(i)}$ for $i = 1, \ldots, k$. By Theorem 1.9, since p_k divides $p'_1 \cdots p'_\ell$, there exists an integer $j_0 \in \{1, \ldots, \ell\}$ such that p_k divides p'_{j_0}, and so $p_k = p'_{j_0}$ since p'_{j_0} is prime. Therefore,

$$\frac{n}{p_k} = p_1 \cdots p_{k-1} = \prod_{\substack{j=1 \\ j \neq j_0}}^{\ell} p'_j < n.$$

It follows from the induction hypothesis that $k - 1 = \ell - 1$, and there is a one-to-one map σ from $\{1, \ldots, k-1\}$ into $\{1, \ldots, k\} \setminus \{j_0\}$ such that $p_i = p'_{\sigma(i)}$ for $i = 1, \ldots, k-1$. Let $\sigma(k) = j_0$. This defines the permutation σ, and the proof is complete. □

For any nonzero integer n and prime number p, we define $v_p(n)$ as the greatest integer r such that p^r divides n. Then $v_p(n)$ is a nonnegative integer, and $v_p(n) \geq 1$ if and only if p divides n. If $v_p(n) = r$, then we say that the prime power p^r *exactly divides* n, and write $p^r \| n$. The *standard factorization* of n is

$$n = \prod_{p|n} p^{v_p(n)}.$$

Since every positive integer is divisible by only a finite number of primes, we can also write

$$n = \prod_{p} p^{v_p(n)},$$

where the product is an infinite product over the set of all prime numbers, and $v_p(n) = 0$ and $p^{v_p(n)} = 1$ for all but finitely many primes p. The function $v_p(n)$ is called the *p-adic value of* n. It is *completely additive* in the sense that $v_p(mn) = v_p(m) + v_p(n)$ for all positive integers m and n (Exercise 13). For example, since $n! = 1 \cdot 2 \cdot 3 \cdots n$, we have

$$v_p(n!) = \sum_{m=1}^{n} v_p(m).$$

The standard factorizations of the first 60 integers are

$1 = 1$	$21 = 3 \cdot 7$	$41 = 41$
$2 = 2$	$22 = 2 \cdot 11$	$42 = 2 \cdot 3 \cdot 7$
$3 = 3$	$23 = 23$	$43 = 43$
$4 = 2^2$	$24 = 2^3 \cdot 3$	$44 = 2^2 \cdot 11$
$5 = 5$	$25 = 5^2$	$45 = 3^2 \cdot 5$
$6 = 2 \cdot 3$	$26 = 2 \cdot 13$	$46 = 2 \cdot 23$
$7 = 7$	$27 = 3^3$	$47 = 47$
$8 = 2^3$	$28 = 2^2 \cdot 7$	$48 = 2^4 \cdot 3$
$9 = 3^2$	$29 = 29$	$49 = 7^2$
$10 = 2 \cdot 5$	$30 = 2 \cdot 3 \cdot 5$	$50 = 2 \cdot 5^2$
$11 = 11$	$31 = 31$	$51 = 3 \cdot 17$
$12 = 2^2 \cdot 3$	$32 = 2^5$	$52 = 2^2 \cdot 13$
$13 = 13$	$33 = 3 \cdot 11$	$53 = 53$
$14 = 2 \cdot 7$	$34 = 2 \cdot 17$	$54 = 2 \cdot 3^3$
$15 = 3 \cdot 5$	$35 = 5 \cdot 7$	$55 = 5 \cdot 11$
$16 = 2^4$	$36 = 2^2 \cdot 3^2$	$56 = 2^3 \cdot 7$
$17 = 17$	$37 = 37$	$57 = 3 \cdot 19$
$18 = 2 \cdot 3^2$	$38 = 2 \cdot 19$	$58 = 2 \cdot 29$
$19 = 19$	$39 = 3 \cdot 13$	$59 = 59$
$20 = 2^2 \cdot 5$	$40 = 2^3 \cdot 5$	$60 = 2^2 \cdot 3 \cdot 5.$

Let $a_1, \ldots, a_k$ be nonzero integers. An integer m' is called a *common multiple* of $a_1, \ldots, a_k$ if it is a multiple of a_i for all $i = 1, \ldots, k$, that is, every integer a_i divides m'. The *least common multiple* of $a_1, \ldots, a_k$ is a positive integer m such that m is a common multiple of $a_1, \ldots, a_k$, and m divides every common multiple of $a_1, \ldots, a_k$. For example, 910 is a common multiple of 35 and 91, and 455 is the least common multiple. We shall show that there is a unique least common multiple for every finite set of nonzero integers. We denote by $[a_1, \ldots, a_k]$ the least common multiple of $a_1, \ldots, a_k$.

Theorem 1.11 *Let $a_1, \ldots, a_k$ be positive integers. Then*

$$(a_1, \ldots, a_k) = \prod_p p^{\min\{v_p(a_1), \ldots, v_p(a_k)\}}$$

and

$$[a_1, \ldots, a_k] = \prod_p p^{\max\{v_p(a_1), \ldots, v_p(a_k)\}}.$$

Proof. This follows immediately from the fundamental theorem of arithmetic. □

Let x be a real number. Recall that the *integer part* of x is the greatest integer not exceeding x, that is, the unique integer n such that $n \leq x <$

$n+1$. We denote the integer part of x by $[x]$. For example, $\left[\frac{4}{3}\right] = 1$, $[\sqrt{7}] = 2$, and $\left[-\frac{4}{3}\right] = -2$. The *fractional part* of x is the real number

$$\{x\} = x - [x] \in [0, 1).$$

Thus, $\left\{\frac{4}{3}\right\} = \frac{1}{3}$ and $\left\{-\frac{4}{3}\right\} = \frac{2}{3}$. We can use the greatest integer function to compute the standard factorization of factorials.

Theorem 1.12 *For every positive integer n and prime p,*

$$v_p(n!) = \sum_{r=1}^{\left[\frac{\log n}{\log p}\right]} \left[\frac{n}{p^r}\right].$$

Proof. Let $1 \leq m \leq n$. If p^r divides m, then $p^r \leq m \leq n$ and $r \leq \log n / \log p$. Since r is an integer, we have $r \leq [\log n / \log p]$ and

$$v_p(m) = \sum_{\substack{r=1 \\ p^r \mid m}}^{\left[\frac{\log n}{\log p}\right]} 1.$$

The number of positive integers not exceeding n that are divisible by p^r is exactly $[n/p^r]$, and so

$$\begin{aligned} v_p(n!) &= \sum_{m=1}^{n} v_p(m) = \sum_{m=1}^{n} \sum_{\substack{r=1 \\ p^r \mid m}}^{\left[\frac{\log n}{\log p}\right]} 1 \\ &= \sum_{r=1}^{\left[\frac{\log n}{\log p}\right]} \sum_{\substack{m=1 \\ p^r \mid m}}^{n} 1 = \sum_{r=1}^{\left[\frac{\log n}{\log p}\right]} \left[\frac{n}{p^r}\right]. \end{aligned}$$

This completes the proof. □

We shall use Theorem 1.12 to compute the standard factorization of 10!. The primes not exceeding 10 are $2, 3, 5$, and 7, and

$$v_2(10!) = \left[\frac{10}{2}\right] + \left[\frac{10}{4}\right] + \left[\frac{10}{8}\right] = 5 + 2 + 1 = 8,$$

$$v_3(10!) = \left[\frac{10}{3}\right] + \left[\frac{10}{9}\right] = 4,$$

$$v_5(10!) = \left[\frac{10}{5}\right] = 2,$$

$$v_7(10!) = \left[\frac{10}{7}\right] = 1.$$

Therefore,

$$10! = 2^8 3^4 5^2 7.$$

For every nonzero integer m, the *radical of* m, denoted by $\mathrm{rad}(m)$, is the product of the distinct primes that divide m, that is,

$$\mathrm{rad}(m) = \prod_{p \mid m} p = \prod_{v_p(m) \geq 1} p.$$

For example, $\mathrm{rad}(15) = \mathrm{rad}(-45) = \mathrm{rad}(225) = 15$ and $\mathrm{rad}(p^r) = p$ for p prime and $r \geq 1$.

Theorem 1.13 *Let m and a be nonzero integers. There exists a positive integer k such that m divides a^k if and only if $\mathrm{rad}(m)$ divides $\mathrm{rad}(a)$.*

Proof. We know that m divides a^k if and only if $v_p(m) \leq v_p(a^k) = kv_p(a)$ for every prime p (Exercise 14). If there exists an integer k such that m divides a^k, then $v_p(a) > 0$ whenever $v_p(m) > 0$, and so every prime that divides m also divides a. This implies that $\mathrm{rad}(m)$ divides $\mathrm{rad}(a)$.

Conversely, if $\mathrm{rad}(m)$ divides $\mathrm{rad}(a)$, then $v_p(a) > 0$ for every prime p such that $v_p(m) > 0$. Since only finitely many primes divide m, it follows that there exists a positive integer k such that $v_p(a^k) = kv_p(a) \geq v_p(m)$ for all primes p, and so m divides a^k. □

Exercises

1. Factor $51,948$ into a product of primes.

2. Factor $10^k + 1$ into a product of primes for $k = 1, 2, 3, 4, 5$.

3. Find the greatest common divisor and least common multiple of $a = 2^3 3^8 7^{12} 13^2$ and $b = 3^6 5^5 11^2 13$.

4. Compute the least common multiple of the integers $1, 2, 3, \ldots, 15$.

5. Compute the standard factorization of $15!$.

6. Prove that $n, n+2, n+4$ are all primes if and only if $n = 3$.

7. Prove that $n, n+4, n+8$ are all primes if and only if $n = 3$.

8. Let $n \geq 2$. Prove that $(n+1)! + k$ is composite for $k = 2, \ldots, n+1$. This shows that there exist arbitrarily long intervals of composite numbers.

9. Prove that $n^5 - n$ is divisible by 30 for every integer n.

10. Find all primes p such that $29p + 1$ is a square.

11. The prime numbers p and q are called *twin primes* if $|p - q| = 2$. Let p and q be primes. Prove that $pq + 1$ is a square if and only if p and q are twin primes.

12. Prove that if p and q are twin primes greater than 3, then $p + q$ is divisible by 12.

13. Let m, n, and k be positive integers. Prove that

$$v_p(mn) = v_p(m) + v_p(n) \qquad \text{and} \qquad v_p(m^k) = kv_p(m).$$

14. Let d and m be nonzero integers. Prove that d divides m if and only if $v_p(d) \leq v_p(m)$ for all primes p.

15. Let $m = \prod_{i=1}^{k} p_i^{r_i}$, where $p_1, \ldots, p_k$ are distinct primes, $k \geq 2$, and $r_i \geq 1$ for $i = 1, \ldots, k$. Let $m_i = mp_i^{-k_i}$ for $i = 1, \ldots, k$. Prove that $(m_1, \ldots, m_k) = 1$.

16. Let a, b, and c be positive integers. Prove that $(ab, c) = 1$ if and only if $(a, c) = (b, c) = 1$.

17. Prove that if 6 divides m, then there exist integers b and c such that $m = bc$ and 6 divides neither b nor c.

18. Prove the following statement or construct a counterexample: If d is composite and d divides m, then there exist integers b and c such that $m = bc$ and d divides neither b nor c.

19. Let a and b be positive integers. Prove that $(a, bc) = (a, b)(a, c)$ for every positive integer c if and only if $(a, b) = 1$.

20. Let $m_1, \ldots, m_k$ be pairwise relatively prime positive integers, and let d divide $m_1 \cdots m_k$. Prove that for each $i = 1, \ldots, k$ there exists a unique divisor d_i of m_i such that $d = d_1 \cdots d_k$.

21. Let $n \geq 2$. Prove that the equation $y^n = 2x^n$ has no solution in positive integers.

22. Let $n \geq 2$, and let x be a rational number. Prove that $\sqrt[n]{x}$ is rational if and only if $x = y^n$ for some rational number y.

23. Let $m_1, \ldots, m_k$ be positive integers and $m = [m_1, \ldots, m_k]$. Prove that there exist positive integers $d_1, \ldots, d_k$ such that d_i is a divisor of m_i for $i = 1, \ldots, k$, $(d_i, d_j) = 1$ for $1 \leq i < j \leq n$, and $m = [d_1, \ldots, d_k] = d_1 \cdots d_k$.

24. Prove that for any positive integers a and b,

$$[a, b] = \frac{ab}{(a, b)}.$$

25. Let a and b be positive integers with $(a,b) = d$. Prove that
$$\left[\frac{a}{d}, \frac{b}{d}\right] = \frac{[a,b]}{d}.$$

26. Prove that for any positive integers a, b, c,
$$[a,b,c] = \frac{abc(a,b,c)}{(a,b)(b,c)(c,a)}.$$

27. Let $a_1, \ldots, a_k$ be positive integers. Prove that $[a_1, \ldots, a_k] = a_1 \cdots a_k$ if and only if the integers $a_1, \ldots, a_k$ are pairwise relatively prime.

28. Let a and b be positive integers and p a prime. Prove that if p divides $[a,b]$ and p divides $a+b$, then p divides (a,b).

29. Let a and b be positive integers such that
$$a + b = 57$$
and
$$[a,b] = 680.$$
Find a and b.

Hint: Show that a and b are relatively prime. Then $a(57-a) = ab = [a,b]$.

30. Let $a\mathbf{Z} = \{ax : x \in \mathbf{Z}\}$ denote the set of all multiples of a. Prove that for any integers $a_1, \ldots, a_k$,
$$\bigcap_{i=1}^{k} a_i\mathbf{Z} = [a_1, \ldots, a_k]\mathbf{Z}.$$

31. A positive integer is called *square-free* if it is the product of distinct prime numbers. Prove that every positive integer can be written uniquely as the product of a square and a square-free integer.

32. Prove that the set of all rational numbers of the form a/b, where $a, b \in \mathbf{Z}$ and b is square-free, is an additive subgroup of $\mathbf{Q}$.

33. A *powerful number* is a positive integer n such that if a prime p divides n, then p^2 divides n. Prove that every powerful number can be written as the product of a square and a cube. Construct examples to show that this representation of powerful numbers is not unique.

34. Prove that m is square-free if and only if $\mathrm{rad}(m) = m$.

35. Prove that $\mathrm{rad}(mn) = \mathrm{rad}(m)\mathrm{rad}(n)$ if and only if $(m,n) = 1$.

36. Let $H = \{1, 5, 9, \ldots\}$ be the arithmetic progression of all positive integers of the form $4k+1$. Elements of H are called *Hilbert numbers.* Show that H is closed under multiplication, that is, $x, y \in H$ implies $xy \in H$. An element x of H will be called a Hilbert prime if $x \neq 1$ and x cannot be written as the product of two strictly smaller elements of H. Compute all the Hilbert primes up to 100. Prove that every element of H can be factored into a product of Hilbert primes, but that unique factorization does not hold in H.

   Hint: Find two essentially distinct factorizations of 441 into a product of Hilbert primes.

37. For $n \geq 1$, consider the rational number

$$h_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

   Prove that h_n is not an integer for any $n \geq 2$.

   Hint: Let 2^a be the largest power of 2 not exceeding n. Let P be the product of the odd positive integers not exceeding n. Consider the number $2^{a-1}Ph_n$.

1.5 Euclid's Theorem and the Sieve of Eratosthenes

How many primes are there? The fundamental theorem of arithmetic tells us that every number is uniquely the product of primes, but it does not give us the number of primes. Euclid proved that the number of primes is infinite. The following proof is also due to Euclid. It has retained its power for more than two thousand years.

Theorem 1.14 (Euclid's theorem) *There are infinitely many primes.*

Proof. Let $p_1, \ldots, p_n$ be any finite set of prime numbers. Consider the integer

$$N = p_1 \cdots p_n + 1.$$

Since $N > 1$, it follows from the fundamental theorem of arithmetic that N is divisible by some prime p. If $p = p_i$ for some $i = 1, \ldots, n$, then p divides $N - p_1 \cdots p_n = 1$, which is absurd. Therefore, $p \neq p_i$ for all $i = 1, \ldots, n$. This means that, for any finite set of primes, there always exists a prime that does not belong to the set, and so the number of primes is infinite. □

Let $\pi(x)$ denote the number of primes not exceeding x. Then $\pi(x) = 0$ for $x < 2$, $\pi(x) = 1$ for $2 \leq x < 3$, $\pi(x) = 2$ for $3 \leq x < 5$, and so on.

Euclid's theorem says that there are infinitely many prime numbers, that is,

$$\lim_{x\to\infty} \pi(x) = \infty,$$

but it does not tell us how to determine them. We can compute all the prime numbers up to x by using a beautiful and efficient method called the *sieve of Eratosthenes*. The sieve is based on a simple observation. If the positive integer n is composite, then n can be written in the form $n = dd'$, where $1 < d \leq d' < n$. If $d > \sqrt{n}$, then

$$n = dd' > \sqrt{n}\sqrt{n} = n,$$

which is absurd. Therefore, if n is composite, then n has a divisor d such that $1 < d \leq \sqrt{n}$. In particular, every composite number $n \leq x$ is divisible by a prime $p \leq \sqrt{x}$.

To find all the primes up to x, we write down the integers between 1 and x, and eliminate numbers from the list according to the following rule: Cross out 1. The first number in the list that is not eliminated is 2; cross out all multiples of 2 that are greater than 2. The iterative procedure is as follows: Let d be the smallest number on the list whose multiples have not already been eliminated. If $d \leq \sqrt{x}$, then cross out all multiples of d that are greater than d. If $d > \sqrt{x}$, stop. This algorithm must terminate after at most $\sqrt{x}$ steps. The prime numbers up to x are the numbers that have not been crossed out.

We shall demonstrate this method to find the prime numbers up to 60. We must sieve out by the prime numbers less than $\sqrt{60}$, that is, by $2, 3, 5$, and 7. Here is the list of numbers up to 60:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60

We cross out 1 and all multiples of 2 beginning with 4:

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	9	~~10~~
11	~~12~~	13	~~14~~	15	~~16~~	17	~~18~~	19	~~20~~
21	~~22~~	23	~~24~~	25	~~26~~	27	~~28~~	29	~~30~~
31	~~32~~	33	~~34~~	35	~~36~~	37	~~38~~	39	~~40~~
41	~~42~~	43	~~44~~	45	~~46~~	47	~~48~~	49	~~50~~
51	~~52~~	53	~~54~~	55	~~56~~	57	~~58~~	59	~~60~~

Next we cross out all multiples of 3 beginning with 6:

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	25	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	35	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	49	~~50~~
~~51~~	~~52~~	53	~~54~~	55	~~56~~	~~57~~	~~58~~	59	~~60~~

Next we cross out all multiples of 5 beginning with 10:

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	49	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~

Finally, we cross out all multiples of 7 beginning with 14:

~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~

The numbers that have not been crossed out are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59.$$

These are the prime numbers up to 60.

Exercises

1. Use the sieve of Eratosthenes to find the prime numbers up to 210. Compute $\pi(210)$.

2. Let $N = 210$. Prove that $N - p$ is prime for every prime p such that $N/2 < p < N$. Find a prime number $q < N/2$ such that $N - q$ is composite.

3. Let $N = 105$. Show that $N - 2^n$ is prime whenever $2 \leq 2^n < N$. This statement is also true for $N = 7, 15, 21, 45$, and 75. It is not known whether $N = 105$ is the largest integer with this property.

4. Let $N = 199$. Show that $N - 2n^2$ is prime whenever $2n^2 < N$. It is not known whether $N = 199$ is the largest integer with this property.

5. Let a and n be positive integers. Prove that $a^n - 1$ is prime only if $a = 2$ and $n = p$ is prime. Primes of the form $M_p = 2^p - 1$ are called *Mersenne primes*. Compute the first five Mersenne primes. The largest known primes are Mersenne primes. It is an unsolved problem to determine whether there are infinitely many Mersenne primes. There is a list of all known Mersenne primes in the Notes at the end of this chapter.

6. Let k be a positive integer. Prove that if $2^k + 1$ is prime, then $k = 2^n$. The integer
$$F_n = 2^{2^n} + 1$$
is called the nth *Fermat number*. Primes of the form $2^{2^n} + 1$ are called *Fermat primes*. Show that F_n is prime for $n = 1, 2, 3, 4$.

7. Prove that F_5 is divisible by 641, and so F_5 is composite.

   Hint: Observe that
$$F_5 = 2^{2^5} - 1 = (2^{32} + 5^4 \cdot 2^{28}) - (5^4 \cdot 2^{28} - 1)$$
and
$$641 = 2^4 + 5^4 = 5 \cdot 2^7 + 1.$$
Prove that 641 divides both $5^4 \cdot 2^{28} + 2^{32}$ and $5^4 \cdot 2^{28} - 1$.

   It is an unsolved problem to determine whether there are infinitely many Fermat primes. Indeed, we do not know whether F_n is prime for any $n > 4$.

8. Modify the proof of Theorem 1.14 to prove that there are infinitely many prime numbers whose remainder is 3 when divided by 4.

   Hint: Let $p_1, p_2, \ldots, p_n$ be primes of the form $4k + 3$, $p_i \neq 3$. Let $N = 4p_1p_2 \cdots p_n + 3$. Show that N must be divisible by some prime q of the form $4k + 3$.

9. Show that every prime number except 2 and 3 has a remainder of 1 or 5 when divided by 6. Prove that there are infinitely many prime numbers whose remainder is 5 when divided by 6.

10. Prove that $\pi(n) \leq n/2$ for $n \geq 8$.

11. Prove that $\pi(n) \leq n/3$ for $n \geq 33$.

    Hint: Prove the following assertions. (i) If $n_0 \geq 3$, then there are at most two primes among the 6 consecutive integers $n_0 + 1, n_0 + 2, \ldots, n_0 + 6$. (ii) Suppose that $n_0 \geq 3$ and $\pi(n_0) \leq n_0/3$. Let $n = n_0 + 6k$ for some positive integer k. Then $\pi(n) \leq n/3$. (iii) Show (by computation) that $\pi(32) > 32/3$ but $\pi(n_0) \leq n_0/3$ for $n_0 = 33, 34, \ldots, 38$. (iv) Show that every integer $n \geq 33$ can be written in the form $n_0 + 6k$ for some nonnegative integer k and $n_0 \in \{33, 34, \ldots, 38\}$.

12. Let $n_0 \geq 6$. Prove that if $\pi(n_0) \leq 4n_0/15$ and $n = n_0 + 30k$, then $\pi(n) \leq 4n/15$.

13. Let $2 = p_1 < p_2 < \cdots$ be the sequence of primes in increasing order. Prove that
$$p_n \leq 2^{2^{n-1}}$$
for all $n \geq 1$.

    Hint: Show that the method used to prove Euclid's theorem (Theorem 1.14) also proves that $p_{n+1} \leq p_1 \cdots p_n + 1$.

14. Let $\log_2 x$ denote the logarithm of x to the base 2. Prove that
$$\pi(x) > \log_2 \log_2 x$$
for all $x > 1$.

    Hint: Exercise 14.

15. Let $p_1, \ldots, p_k$ be a finite set of prime numbers. Prove that the number of positive integers $n \leq x$ that can be written in the form $n = p_1^{r_1} \cdots p_k^{r_k}$ is at most
$$\prod_{i=1}^{k} \left(\frac{\log x}{\log p_i} + 1 \right).$$
Prove that if x is sufficiently large, then there are positive integers $n \leq x$ that cannot be represented in this way. Use this to give another proof that the number of primes is infinite.

1.6 A Linear Diophantine Equation

A *diophantine equation* is an equation of the form
$$f(x_1, \ldots, x_k) = b$$
that we want to solve in rational numbers, integers, or nonnegative integers. This means that the values of the variables $x_1, \ldots, x_k$ will be rationals, integers, or nonnegative integers. Usually the function $f(x_1, \ldots, x_k)$ is a polynomial with rational or integer coefficients.

In this section we consider the linear diophantine equation
$$a_1 x_1 + \cdots + a_k x_k = b.$$
We want to know when this equation has a solution in integers, and when it has a solution in nonnegative integers. For example, the equation
$$3x_1 + 5x_2 = b$$

has a solution in integers for every integer b, and a solution in nonnegative integers for $b = 0, 3, 5, 6$, and all $b \geq 8$ (Exercise 1).

Theorem 1.15 *Let $a_1, \ldots, a_k$ be integers, not all zero. For any integer b, there exist integers $x_1, \ldots, x_k$ such that*

$$a_1x_1 + \cdots + a_kx_k = b \tag{1.4}$$

if and only if b is a multiple of $(a_1, \ldots, a_k)$. In particular, the linear equation (1.4) has a solution for every integer b if and only if the numbers $a_1, \ldots, a_k$ are relatively prime.

Proof. Let $d = (a_1, \ldots, a_k)$. If equation (1.4) is solvable in integers x_i, then d divides b since d divides each integer a_i. Conversely, if d divides b, then $b = dq$ for some integer q. By Theorem 1.4, there exist integers $y_1, \ldots, y_k$ such that

$$a_1y_1 + \cdots + a_ky_k = d.$$

Let $x_i = y_iq$ for $i = 1, \ldots, k$. Then

$$a_1x_1 + \cdots + a_kx_k = a_1(y_1q) + \cdots + a_k(y_kq) = dq = b$$

is a solution of (1.4). It follows that (1.4) is solvable in integers for every b if and only if $(a_1, \ldots, a_k) = 1$. □

Theorem 1.16 *Let $a_1, \ldots, a_k$ be positive integers such that*

$$(a_1, \ldots, a_k) = 1.$$

If

$$b \geq (a_k - 1)\sum_{i=1}^{k-1} a_i,$$

then there exist nonnegative integers $x_1, \ldots, x_k$ such that

$$a_1x_1 + \cdots + a_kx_k = b.$$

Proof. By Theorem 1.15, there exist integers $z_1, \ldots, z_k$ such that

$$a_1z_1 + \cdots + a_kz_k = b.$$

Using the division algorithm, we can divide each of the integers $z_1, \ldots, z_{k-1}$ by a_k so that

$$z_i = a_kq_i + x_i$$

and

$$0 \leq x_i \leq a_k - 1$$

for $i = 1, \ldots, k-1$. Let

$$x_k = z_k + \sum_{i=1}^{k-1} a_i q_i.$$

Then

$$\begin{aligned} b &= a_1 z_1 + \cdots + a_{k-1} z_{k-1} + a_k z_k \\ &= a_1(a_k q_1 + x_1) + \cdots + a_{k-1}(a_k q_{k-1} + x_{k-1}) + a_k z_k \\ &= a_1 x_1 + \cdots + a_{k-1} x_{k-1} + a_k \left(z_k + \sum_{i=1}^{k-1} a_i q_i \right) \\ &= a_1 x_1 + \cdots + a_{k-1} x_{k-1} + a_k x_k \\ &\leq (a_k - 1) \sum_{i=1}^{k-1} a_i + a_k x_k, \end{aligned}$$

where x_k is an integer, possibly negative. However, if

$$b \geq (a_k - 1) \sum_{i=1}^{k-1} a_i,$$

then $a_k x_k \geq 0$ and so $x_k \geq 0$. This completes the proof. □

Let $a_1, \ldots, a_k$ be relatively prime positive integers. Since every sufficiently large integer can be written as a nonnegative integral linear combination of $a_1, \ldots, a_k$, it follows that there exists a smallest integer

$$G(a_1, \ldots, a_k)$$

such that every integer $b \geq G(a_1, \ldots, a_k)$ can be represented in the form (1.4), where the variables $x_1, \ldots, x_k$ are nonnegative integers. The example above shows that

$$G(3, 5) = 8.$$

The *linear diophantine problem of Frobenius* is to determine $G(a_1, \ldots, a_k)$ for all finite sets of relatively prime positive integers $a_1, \ldots, a_k$. This is a difficult open problem, but there are some special cases where the solution is known. The following theorem solves the Frobenius problem in the case $k = 2$.

Theorem 1.17 *Let a_1 and a_2 be relatively prime positive integers. Then*

$$G(a_1, a_2) = (a_1 - 1)(a_2 - 1).$$

Proof. We saw in the proof of Theorem 1.15 that for every integer b there exist integers x_1 and x_2 such that

$$b = a_1x_1 + a_2x_2 \qquad \text{and} \qquad 0 \leq x_1 \leq a_2 - 1. \tag{1.5}$$

If we have another representation

$$b = a_1x_1' + a_2x_2', \qquad \text{and} \qquad 0 \leq x_1' \leq a_2 - 1,$$

then

$$a_1(x_1 - x_1') = a_2(x_2' - x_2).$$

Since a_2 divides $a_1(x_1 - x_1')$ and $(a_1, a_2) = 1$, Euclid's lemma (Theorem 1.7) implies that a_2 divides $x_1 - x_1'$. Then $x_1 = x_1'$, since $|x_1 - x_1'| \leq a_2 - 1$. It follows that $x_2 = x_2'$, and so the representation (1.5) is unique.

If the integer b *cannot* be represented as a nonnegative integral combination of a_1 and a_2, then we must have $x_1 \leq -1$ in the representation (1.5). This implies that

$$b = a_1x_1 + a_2x_2 \leq a_1(a_2 - 1) + a_2(-1) = (a_1 - 1)(a_2 - 1) - 1,$$

and so $G(a_1, a_2) \leq (a_1 - 1)(a_2 - 1)$. On the other hand, since

$$a_1(a_2 - 1) + a_2(-1) = a_1a_2 - a_1 - a_2 < a_1a_2,$$

it follows that if

$$a_1a_2 - a_1 - a_2 = a_1x_1 + a_2x_2$$

for any nonnegative integers x_1 and x_2, then $0 \leq x_1 \leq a_2 - 1$. By the uniqueness of the representation (1.5), we must have $x_1 = a_2 - 1$ and $x_2 = -1$. Therefore, the integer $a_1a_2 - a_1 - a_2$ cannot be represented as a nonnegative integral linear combination of a_1 and a_2, and so $G(a_1, a_2) = (a_1 - 1)(a_2 - 1)$. □

Exercises

1. Prove that the equation

$$3x_1 + 5x_2 = b$$

has a solution in integers for every integer b, and a solution in nonnegative integers for $b = 0, 3, 5, 6$ and all $b \geq 8$.

2. Find all solutions in nonnegative integers x_1 and x_2 of the linear diophantine equation

$$2x_1 + 7x_2 = 53.$$

3. Find all solutions in nonnegative integers x_1 and x_2 of the linear diophantine equation

$$28x_1 + 35x_2 = 136.$$

4. Let a_1 and a_2 be relatively prime positive integers. Let $N(a_1, a_2)$ denote the number of nonnegative integers that cannot be represented in the form

$$a_1x_1 + a_2x_2$$

with x_1, x_2 nonnegative integers. Compute $N(3, 10)$ and $N(3, 10)/G(3, 10)$.

5. Compute $N(7, 8)$ and $N(7, 8)/G(7, 8)$.

6. Find all nonnegative integers that cannot be represented by the form

$$3x_1 + 10x_2 + 14x_3$$

with x_1, x_2, x_3 nonnegative integers. Compute $G(3, 10, 14)$.

7. Let a_1 and a_2 be relatively prime positive integers. Let $\mathcal{M}$ be the set of all integers n such that $0 \leq n \leq a_1a_2 - a_1 - a_2$ and n can be written in the form $n = a_1x_1 + a_2x_2$, where x_1 and x_2 are nonnegative integers. Let $\mathcal{N}$ be the set of all integers n such that $0 \leq n \leq a_1a_2 - a_1 - a_2$ and n cannot be written in the form $n = a_1x_1 + a_2x_2$, where x_1 and x_2 are nonnegative integers. Then $|\mathcal{N}| = N(a_1, a_2)$ and $|\mathcal{M}| + |\mathcal{N}| = (a_1 - 1)(a_2 - 1)$. Let $n \in [0, a_1a_2 - a_1 - a_2]$, and write n in the form

$$n = a_1x_1 + a_2x_2, \quad \text{where } 0 \leq x_1 \leq a_2 - 1.$$

This representation is unique. Define the function f by

$$f(n) = a_1a_2 - a_1 - a_2 - n = a_1(a_2 - 1 - x_1) - a_2(x_2 + 1).$$

Prove that f is an involution that maps $\mathcal{M}$ onto $\mathcal{N}$ and $\mathcal{N}$ onto $\mathcal{M}$, and so

$$|\mathcal{M}| = |\mathcal{N}| = \frac{(a_1 - 1)(a_2 - 1)}{2}$$

and

$$\frac{N(a_1, a_2)}{G(a_1, a_2)} = \frac{1}{2}.$$

8. Find all solutions in nonnegative integers x_1, x_2, and x_3 of the linear diophantine equation

$$6x_1 + 10x_2 + 15x_3 = 30.$$

9. Find all solutions in integers x_1, x_2, and x_3 of the system of linear diophantine equations

$$3x_1 + 5x_2 + 7x_3 = 560,$$

$$9x_1 + 25x_2 + 49x_3 = 2920.$$

10. Find all solutions of the *Ramanujan-Nagell diophantine equation*

$$x^2 + 7 = 2^n$$

with $x \leq 1000$.

11. Find all solutions of the *Ljunggren diophantine equation*

$$x^2 - 2y^4 = -1$$

with $x \leq 1000$.

12. When is the sum of a geometric progression equal to a power? Equivalently, what are the solutions of the exponential diophantine equation

$$1 + x + x^2 + \cdots + x^m = y^n \tag{1.6}$$

in integers x, m, y, n greater than 2? Check that

$$1 + 3 + 3^2 + 3^3 + 3^4 = 11^2,$$

$$1 + 7 + 7^2 + 7^3 = 20^2,$$

and

$$1 + 18 + 18^2 = 7^3.$$

These are the only known solutions of (1.6).

1.7 Notes

I can hardly do better than go back to the Greeks. I will state and prove two of the famous theorems of Greek mathematics. They are 'simple' theorems, simple both in idea and in execution, but there is no doubt at all about their being theorems of the highest class. Each is as fresh and significant as when it was discovered—two thousand years have not written a wrinkle on either of them.

G. H. Hardy [51, p. 92]

Number theory is an ancient subject. The famous theorems to which Hardy refers are the theorems that there are infinitely many primes (Theorem 1.14) and that $\sqrt{2}$ is irrational (Exercise 22 in Section 1.4). These appear in Euclid's *Elements* [61, Book IX, Proposition 20, and Book X, Proposition 9]. The Euclidean algorithm also appears in Euclid [61, Book VII, Proposition 2]. For fragments of number theory in Babylonian mathematics, see Neugebauer [110] and van der Waerden [147].

There are many excellent introductions to elementary number theory. My favorite is *Number Theory for Beginners* by André Weil [152]. Two classic works are Hardy and Wright [60] and Landau [87]. Other interesting books are Davenport [22], Hua [68], Kumanduri and Romero [85] and Ireland and Rosen [72]. There are beautiful introductions to algebraic number theory by Borevich and Shafarevich [13], Hecke [63, 64], Lang [90], and Neukirch [111], and to analytic number theory by Apostol [3], Davenport [21], Rademacher [119], and Serre [131, 132]. An excellent survey volume is Manin and Panchishkin, *Introduction to Number Theory* [96].

The best history is Weil, *Number Theory: An Approach through History. From Hammurapi to Legendre* [153]. There is also Leonard Eugene Dickson's encyclopedic but unreadable three-volume *History of the Theory of Numbers* [25].

Guy's *Unsolved Problems in Number Theory* [45] is a nice survey of unusual problems and results in elementary number theory.

For a refinement of Theorem 1.16, see Nathanson [101].

Lang's *Algebra* [89] is the standard reference for the algebra used in this book.

In October, 1999, only 38 Mersenne primes had been discovered. The list of these primes is as follows:

2^2-1	2^3-1	2^5-1	2^7-1	$2^{13}-1$
$2^{17}-1$	$2^{19}-1$	$2^{31}-1$	$2^{61}-1$	$2^{89}-1$
$2^{107}-1$	$2^{127}-1$	$2^{521}-1$	$2^{607}-1$	$2^{1279}-1$
$2^{2203}-1$	$2^{2281}-1$	$2^{3217}-1$	$2^{4253}-1$	$2^{4423}-1$
$2^{9689}-1$	$2^{9941}-1$	$2^{11213}-1$	$2^{19937}-1$	$2^{21701}-1$
$2^{23209}-1$	$2^{44497}-1$	$2^{86243}-1$	$2^{110503}-1$	$2^{132049}-1$
$2^{216091}-1$	$2^{756839}-1$	$2^{859433}-1$	$2^{1257787}-1$	$2^{1398269}-1$
$2^{2976221}-1$	$2^{3021377}-1$	$2^{6972593}-1.$		

The largest prime known in October, 1999 was the Mersenne prime $M_{6972593}$. An Internet site devoted to Mersenne primes and related problems in number theory is www.mersenne.org.

2
Congruences

2.1 The Ring of Congruence Classes

Let m be a positive integer. If a and b are integers such that $a-b$ is divisible by m, then we say that a and b are *congruent modulo* m, and write

$$a \equiv b \pmod{m}.$$

Integers a and b are called *incongruent modulo* m if they are not congruent modulo m. For example, $-12 \equiv 43 \pmod{5}$ and $-12 \equiv 43 \pmod{11}$, but $-12 \not\equiv 43 \pmod{7}$. Every even integer is congruent to 0 modulo 2, and every odd integer is congruent to 1 modulo 2. If x is not divisible by 3, then $x^2 \equiv 1 \pmod{3}$.

Congruence modulo m is an equivalence relation, since for all integers a, b, and c we have

(i) Reflexivity: $a \equiv a \pmod{m}$,

(ii) Symmetry: If $a \equiv b \pmod{m}$, then $b \equiv a \pmod{m}$, and

(iii) Transitivity: If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then $a \equiv c \pmod{m}$.

Properties (i) and (ii) follow immediately from the definition of congruence. To prove (iii), we observe that if $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then there exist integers x and y such that $a - b = mx$ and $b - c = my$. Since

$$a - c = (a - b) + (b - c) = mx + my = m(x + y),$$

it follows that $a \equiv c \pmod{m}$. The equivalence class of an integer a under this relation is called the *congruence class* of a modulo m, and written $a+m\mathbf{Z}$. Thus, $a+m\mathbf{Z}$ is the set of all integers b such that $b \equiv a \pmod{m}$, that is, the set of all integers of the form $a + mx$ for some integer x. If $(a + m\mathbf{Z}) \cap (b + m\mathbf{Z}) \neq \emptyset$, then $a + m\mathbf{Z} = b + m\mathbf{Z}$. We denote by $\mathbf{Z}/m\mathbf{Z}$ the set of all congruence classes modulo m.

A congruence class modulo m is also called a *residue class* modulo m.

By the division algorithm, we can write every integer a in the form $a = mq + r$, where q and r are integers and $0 \leq r \leq m - 1$. Then $a \equiv r \pmod{m}$, and r is called the *least nonnegative residue* of a modulo m.

If $a \equiv 0 \pmod{m}$ and $|a| < m$, then $a = 0$, since 0 is the only integral multiple of m in the open interval $(-m, m)$. This implies that if $a \equiv b \pmod{m}$ and $|a-b| < m$, then $a = b$. In particular, if $r_1, r_2 \in \{0, 1, \ldots, m-1\}$ and if $a \equiv r_1 \pmod{m}$ and $a \equiv r_2 \pmod{m}$, then $r_1 = r_2$. Thus, every integer belongs to a unique congruence class of the form $r + m\mathbf{Z}$, where $0 \leq r \leq m - 1$, and so

$$\mathbf{Z}/m\mathbf{Z} = \{m\mathbf{Z}, 1 + m\mathbf{Z}, \ldots, (m-1) + m\mathbf{Z}\}.$$

The integers $0, 1, \ldots, m-1$ are pairwise incongruent modulo m.

A set of integers $R = \{r_1, \ldots, r_m\}$ is called a *complete set of residues* modulo m if $r_1, \ldots, r_m$ are pairwise incongruent modulo m and every integer x is congruent modulo m to some integer $r_i \in R$. For example, the set $\{0, 2, 4, 6, 8, 10, 12\}$ is a complete set of residues modulo 7. The set $\{0, 3, 6, 9, 12, 15, 18, 21\}$ is a complete set of residues modulo 8. The set $\{0, 1, 2, \ldots, m-1\}$ is a complete set of residues modulo m for every positive integer m.

There is a natural way to define addition, subtraction, and multiplication of congruence classes. If

$$a_1 \equiv a_2 \pmod{m}$$

and

$$b_1 \equiv b_2 \pmod{m},$$

then

$$a_1 + b_1 \equiv a_2 + b_2 \pmod{m},$$

$$a_1 - b_1 \equiv a_2 - b_2 \pmod{m},$$

and

$$a_1 b_1 \equiv a_2 b_2 \pmod{m}.$$

These statements are consequences of the identities

$$(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) \equiv 0 \pmod{m},$$

$$(a_1 - b_1) - (a_2 - b_2) = (a_1 - a_2) - (b_1 - b_2) \equiv 0 \pmod{m}$$

and

$$a_1b_1 - a_2b_2 = a_1(b_1 - b_2) + (a_1 - a_2)b_2 \equiv 0 \pmod{m}.$$

Addition, subtraction, and multiplication in $\mathbf{Z}/m\mathbf{Z}$ are well-defined if we define the sum, difference, and product of congruence classes modulo m by

$$(a + m\mathbf{Z}) + (b + m\mathbf{Z}) = (a + b) + m\mathbf{Z},$$

$$(a + m\mathbf{Z}) - (b + m\mathbf{Z}) = (a - b) + m\mathbf{Z},$$

and

$$(a + m\mathbf{Z}) \cdot (b + m\mathbf{Z}) = ab + m\mathbf{Z}.$$

Addition of congruence classes is associative and commutative, since

$$\begin{aligned} &((a + m\mathbf{Z}) + (b + m\mathbf{Z})) + (c + m\mathbf{Z}) \\ &= ((a + b) + m\mathbf{Z}) + (c + m\mathbf{Z}) \\ &= ((a + b) + c) + m\mathbf{Z} \\ &= (a + (b + c)) + m\mathbf{Z} \\ &= (a + m\mathbf{Z}) + ((b + c) + m\mathbf{Z}) \\ &= (a + m\mathbf{Z}) + ((b + m\mathbf{Z}) + (c + m\mathbf{Z})) \end{aligned}$$

and

$$\begin{aligned} (a + m\mathbf{Z}) + (b + m\mathbf{Z}) &= (a + b) + m\mathbf{Z} \\ &= (b + a) + m\mathbf{Z} \\ &= (b + m\mathbf{Z}) + (a + m\mathbf{Z}). \end{aligned}$$

The congruence class $m\mathbf{Z}$ is a zero element for addition, since $m\mathbf{Z} + (a + m\mathbf{Z}) = a + m\mathbf{Z}$ for all $a + m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$, and the additive inverse of the congruence class $a + m\mathbf{Z}$ is $-a + m\mathbf{Z}$, since

$$(a + m\mathbf{Z}) + (-a + m\mathbf{Z}) = (a - a) + m\mathbf{Z} = m\mathbf{Z}.$$

From these identities we see that the set of congruence classes modulo m is an abelian group under addition.

We have also defined multiplication in $\mathbf{Z}/m\mathbf{Z}$. Multiplication is associative and commutative, since

$$\begin{aligned} ((a + m\mathbf{Z})(b + m\mathbf{Z}))(c + m\mathbf{Z}) &= (ab)c + m\mathbf{Z} \\ &= a(bc) + m\mathbf{Z} \\ &= (a + m\mathbf{Z})((b + m\mathbf{Z})(c + m\mathbf{Z})) \end{aligned}$$

and

$$(a + m\mathbf{Z})(b + m\mathbf{Z}) = ab + m\mathbf{Z} = ba + m\mathbf{Z} = (b + m\mathbf{Z})(a + m\mathbf{Z}).$$

The congruence class $1 + m\mathbf{Z}$ is an identity for multiplication, since

$$(1 + m\mathbf{Z})(a + m\mathbf{Z}) = a + m\mathbf{Z}$$

for all $a + m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$. Finally, multiplication of congruence classes is *distributive with respect to addition* in the sense that

$$\begin{aligned} &(a + m\mathbf{Z})((b + m\mathbf{Z}) + (c + m\mathbf{Z})) \\ &\quad = \quad a(b + c) + m\mathbf{Z}) \\ &\quad = \quad (ab + m\mathbf{Z}) + (ac + m\mathbf{Z}) \\ &\quad = \quad (a + m\mathbf{Z})(b + m\mathbf{Z}) + (a + m\mathbf{Z})(c + m\mathbf{Z}) \end{aligned}$$

for all $a + m\mathbf{Z}, b + m\mathbf{Z}, c + m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$.

A *ring* is a set R with two binary operations, addition and multiplication, such that R is an abelian group under addition with additive identity 0, and multiplication satisfies the following axioms:

(i) Associativity: For all $x, y, z \in R$,

$$(xy)z = x(yz).$$

(ii) Identity element: There exists an element $1 \in R$ such that for all $x \in R$,

$$1 \cdot x = x \cdot 1 = x.$$

The element 1 is called the *multiplicative identity* of the ring.

(iii) Distributivity: For all $x, y, z \in R$,

$$x(y + z) = xy + xz.$$

The ring R is *commutative* if multiplication also satisfies the axiom

(iv) Commutativity: For all $x, y \in R$,

$$xy = yx.$$

The integers, rational numbers, real numbers, and complex numbers are examples of commutative rings. The set $M_2(\mathbf{C})$ of 2×2 matrices with complex coefficients and the usual matrix addition and multiplication is a noncommutative ring.

Let R and S be rings with multiplicative identities 1_R and 1_S, respectively. A map $f : R \to S$ is called a *ring homomorphism* if $f(x + y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$ for all $x, y \in R$, and $f(1_R) = 1_S$.

An element a in the ring R is called a *unit* if there exists an element $x \in R$ such that $ax = xa = 1$. If a is a unit in R and $x \in R$ and $y \in R$ are both inverses of a, then $x = x(ay) = (xa)y = y$, and so the inverse of a is

unique. We denote the inverse of a by a^{-1}. The set $R^{\times}$ of all units in R is a multiplicative group, called the *group of units* in the ring R. A *field* is a commutative ring in which every nonzero element is a unit. For example, the rational, real, and complex numbers are fields. The integers form a ring but not a field, and the only units in the ring of integers are ± 1.

The various properties of sums and products of congruence classes that we proved in this section are equivalent to the following statement.

Theorem 2.1 *For every integer $m \geq 2$, the set $\mathbf{Z}/m\mathbf{Z}$ of congruence classes modulo m is a commutative ring.*

Exercises

1. Compute the least nonnegative residue of $10^k + 1$ modulo 13 for $k = 1, 2, 3, 4$.
2. Compute the least nonnegative residue of 5^{22} modulo 23.
3. Construct the multiplication table for the ring $\mathbf{Z}/5\mathbf{Z}$.
4. Construct the multiplication table for the ring $\mathbf{Z}/6\mathbf{Z}$.
5. Prove that every integer is congruent modulo 9 to one of the even integers $0, 2, 4, 6, \ldots, 16$.
6. Let m be an odd positive integer. Prove that every integer is congruent modulo m to one of the even integers $0, 2, 4, 6, \ldots, 2m-2$.
7. Prove that every integer is congruent modulo 9 to a unique integer r such that $-4 \leq r \leq 4$.
8. Let $m = 2q + 1$ be an odd positive integer. Prove that every integer is congruent modulo m to a unique integer r such that $-q \leq r \leq q$.
9. Let $m = 2q$ be an even positive integer. Prove that every integer is congruent modulo m to a unique integer r such that $-(q-1) \leq r \leq q$.
10. Prove that $a^3 \equiv a \pmod 6$ for every integer a.
11. Prove that $a^4 \equiv 1 \pmod 5$ for every integer a that is not divisible by 5.
12. Prove that if a is an odd integer, then $a^2 \equiv 1 \pmod 8$.
13. Let d be a positive integer that is a common divisor of a, b, and m. Prove that
$$a \equiv b \pmod m$$
if and only if
$$\frac{a}{d} \equiv \frac{b}{d} \pmod{\frac{m}{d}}.$$

14. Prove that if x, y, z are integers such that $x^2 + y^2 = z^2$, then $xyz \equiv 0 \pmod{60}$.

15. Prove that $a_1 \equiv a_2 \pmod{m}$ implies $a_1^k \equiv a_2^k \pmod{m}$ for all $k \geq 1$. Prove that if $f(x)$ is a polynomial with integer coefficients and $a_1 \equiv a_2 \pmod{m}$, then $f(a_1) \equiv f(a_2) \pmod{m}$.

16. (A criterion for divisibility by 9.) Prove that a positive integer n is divisible by 9 if and only if the sum of its decimal digits is divisible by 9. (For example, the sum of the decimal digits of 567 is $5+6+7 = 18$.)

    Hint: Prove that $10^k \equiv 1 \pmod{9}$ for every nonnegative integer k.

17. (A criterion for divisibility by 11.) Prove that a positive integer n is divisible by 11 if and only if the alternating sum of its decimal digits is divisible by 11. (For example, the alternating sum of the decimal digits of 80,729 is $-9 + 2 - 7 + 0 - 8 = -22$.)

    Hint: Prove that $10^k \equiv (-1)^k \pmod{11}$ for every nonnegative integer k.

18. Prove that if $x_1, \ldots, x_m$ is a sequence of m not necessarily distinct integers, then there is a subsequence of consecutive terms whose sum is divisible by m, that is, there exist integers $1 \leq k \leq \ell \leq m$ such that
$$\sum_{i=k}^{\ell} x_i \equiv 0 \pmod{m}.$$

    Hint: Consider the $m+1$ integers $0, x_1, x_1+x_2, x_1+x_2+x_3, \ldots, x_1 + x_2 + \cdots + x_m$.

19. Let $m \geq 2$ and let d be a positive divisor of $m - 1$. Let $n = a_0 + a_1 m + \cdots + a_k m^k$ be the m-adic representation of n. Prove that $n \equiv 0 \pmod{d}$ if and only if $a_0 + a_1 + \cdots + a_k \equiv 0 \pmod{d}$.

20. Let n be a positive integer such that $n \equiv 3 \pmod{4}$. Prove that n cannot be written as the sum of two squares.

21. Prove that every integer belongs to at least one of the following 6 congruence classes:
$$\begin{array}{rl} 0 & \pmod{2} \\ 0 & \pmod{3} \\ 1 & \pmod{4} \\ 3 & \pmod{8} \\ 7 & \pmod{12} \\ 23 & \pmod{24}. \end{array}$$

22. Let p be prime, $m \geq 1$, and $0 \leq k \leq p-1$. Prove that

$$N = \binom{mp+k}{p} \equiv m \pmod{p}.$$

Hint: Consider the integer $(p-1)!N$ modulo p.

23. Let G be the subset of $M_2(\mathbf{C})$ consisting of the four matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Prove that G is a multiplicative group isomorphic to the additive group of congruence classes $\mathbf{Z}/4\mathbf{Z}$.

2.2 Linear Congruences

The following theorem is one of the most useful and important tools in elementary number theory.

Theorem 2.2 *Let m, a, b be integers with $m \geq 1$. Let $d = (a, m)$ be the greatest common divisor of a and m. The congruence*

$$ax \equiv b \pmod{m} \tag{2.1}$$

has a solution if and only if

$$b \equiv 0 \pmod{d}.$$

If $b \equiv 0 \pmod{d}$, then the congruence (2.1) has exactly d solutions in integers that are pairwise incongruent modulo m. In particular, if $(a, m) = 1$, then for every integer b the congruence (2.1) has a unique solution modulo m.

Proof. Let $d = (a, m)$. Congruence (2.1) has a solution if and only if there exist integers x and y such that

$$ax - b = my,$$

or, equivalently,

$$b = ax - my.$$

By Theorem 1.15, this is possible if and only if $b \equiv 0 \pmod{d}$.

If x and x_1 are solutions of (2.1), then

$$a(x_1 - x) \equiv ax_1 - ax \equiv b - b \equiv 0 \pmod{m},$$

and so

$$a(x_1 - x) = mz$$

for some integer z. If d is the greatest common divisor of a and m, then $(a/d, m/d) = 1$ and

$$\left(\frac{a}{d}\right)(x - x_1) = \left(\frac{m}{d}\right) z.$$

By Euclid's lemma (Theorem 1.7), m/d divides $x_1 - x$, and so

$$x_1 = x + \frac{im}{d}$$

for some integer i, that is,

$$x_1 \equiv x \pmod{\frac{m}{d}}.$$

Moreover, every integer x_1 of this form is a solution of (2.1). An integer x_1 congruent to x modulo m/d is congruent to $x + im/d$ modulo m for some integer $i = 0, 1, \ldots d-1$, and the d integers $x+im/d$ with $i = 0, 1, \ldots, d-1$ are pairwise incongruent modulo m. Thus, the congruence (2.1) has exactly d pairwise incongruent solutions. This completes the proof. □

Theorem 2.3 *If p is a prime, then $\mathbf{Z}/p\mathbf{Z}$ is a field.*

Proof. If $a + p\mathbf{Z} \in \mathbf{Z}/p\mathbf{Z}$ and $a + p\mathbf{Z} \neq p\mathbf{Z}$, then a is an integer not divisible by p. By Theorem 2.2, there exists an integer x such that $ax \equiv 1 \pmod{p}$. This implies that

$$(a + p\mathbf{Z})(x + p\mathbf{Z}) = 1 + p\mathbf{Z},$$

and so $a + p\mathbf{Z}$ is invertible. Thus, every nonzero congruence class in $\mathbf{Z}/p\mathbf{Z}$ is a unit and $\mathbf{Z}/p\mathbf{Z}$ is a field. □

Here are some examples of linear congruences. The congruence

$$7x \equiv 3 \pmod{5}$$

has a unique solution modulo 5 since $(7, 5) = 1$. The solution is $x \equiv 4 \pmod{5}$. The congruence

$$35x \equiv -14 \pmod{91} \tag{2.2}$$

is solvable since $(35, 91) = 7$ and

$$-14 \equiv 0 \pmod{7}.$$

Congruence (2.2) is equivalent to the congruence

$$5x \equiv -2 \pmod{13}, \tag{2.3}$$

which has the unique solution $x \equiv 10 \pmod{13}$. Every solution of (2.2) satisfies

$$x \equiv 10 \pmod{13}$$

and so a complete set of solutions that are pairwise incongruent modulo 91 is $\{10, 23, 36, 49, 62, 75, 88\}$.

Lemma 2.1 *Let p be a prime number. Then $x^2 \equiv 1 \pmod{p}$ if and only if $x \equiv \pm 1 \pmod{p}$.*

Proof. If $x \equiv \pm 1 \pmod{p}$, then $x^2 \equiv 1 \pmod{p}$. Conversely, if $x^2 \equiv 1 \pmod{p}$, then p divides $x^2 - 1 = (x-1)(x+1)$, and so p must divide $x - 1$ or $x + 1$. □

Theorem 2.4 (Wilson) *If p is prime, then*

$$(p-1)! \equiv -1 \pmod{p}.$$

Proof. This is true for $p = 2$ and $p = 3$, since $1! \equiv -1 \pmod{2}$ and $2! \equiv -1 \pmod{3}$. Let $p \geq 5$. By Theorem 2.2, to each integer $a \in \{1, 2, \ldots, p-1\}$ there is a unique integer $a^{-1} \in \{1, 2, \ldots, p-1\}$ such that $aa^{-1} \equiv 1 \pmod{p}$. By Lemma 2.1, $a = a^{-1}$ if and only if $a = 1$ or $a = p-1$. Therefore, we can partition the $p-3$ numbers in the set $\{2, 3, \ldots, p-2\}$ into $(p-3)/2$ pairs of integers $\{a_i, a_i^{-1}\}$ such that $a_i a_i^{-1} \equiv 1 \pmod{p}$ for $i = 1, \ldots, (p-3)/2$. Then

$$\begin{aligned} (p-1)! &\equiv 1 \cdot 2 \cdot 3 \cdots (p-2)(p-1) \\ &\equiv (p-1) \prod_{i=1}^{(p-3)/2} a_i a_i^{-1} \\ &\equiv p-1 \\ &\equiv -1 \pmod{p}. \end{aligned}$$

This completes the proof. □

For example,

$$4! \equiv 24 \equiv -1 \pmod{5}$$

and

$$6! \equiv 720 \equiv -1 \pmod{7}.$$

The converse of Wilson's theorem is also true (Exercise 7).

Theorem 2.5 *Let m and d be positive integers such that d divides m. If a is an integer relatively prime to d, then there exists an integer a' such that $a' \equiv a \pmod{d}$ and a' is relatively prime to m.*

Proof. Let $m = \prod_{i=1}^{k} p_i^{r_i}$ and $d = \prod_{i=1}^{k} p_i^{s_i}$, where $r_i \geq 1$ and $0 \leq s_i \leq r_i$ for $i = 1, \ldots, k$. Let m' be the product of the prime powers that divide m but not d. Then

$$m' = \prod_{\substack{i=1 \\ s_i=0}}^{k} p_i^{r_i}$$

and

$$(m', d) = 1.$$

By Theorem 2.2, there exists an integer x such that

$$dx \equiv 1 - a \pmod{m'}.$$

Then

$$a' = a + dx \equiv 1 \pmod{m'}$$

and so

$$(a', m') = 1.$$

Also,

$$a' \equiv a \pmod{d}.$$

If $(a', m) \neq 1$, there exists a prime p that divides both a' and m. However, p does not divide m' since $(a', m') = 1$. It follows that p divides d, and so p divides $a' - dx = a$, which is impossible since $(a, d) = 1$. Therefore, $(a', m) = 1$. □

If $a \equiv b \pmod{m}$, then $a = b + mx$ for some integer x. An integer d is a common divisor of a and m if and only if d is a common divisor of b and m, and so $(a, m) = (b, m)$. In particular, if a is relatively prime to m, then every integer in the congruence class of $a + m\mathbf{Z}$ is relatively prime to m. A congruence class modulo m is called *relatively prime to m* if some (and, consequently, every) integer in the class is relatively prime to m.

We denote by $\varphi(m)$ the number of congruence classes in $\mathbf{Z}/m\mathbf{Z}$ that are relatively prime to m. The function $\varphi(m)$ is called the *Euler phi function.* Equivalently, $\varphi(m)$ is the number of integers in the set $0, 1, 2, \ldots, m - 1$ that are relatively prime to m. The Euler phi function is also called the *totient function.*

A set of integers $\{r_1, \ldots, r_{\varphi(m)}\}$ is called a *reduced set of residues* modulo m if every integer x such that $(x, m) = 1$ is congruent modulo m to some integer r_i. For example, the sets $\{1, 2, 3, 4, 5, 6\}$ and $\{2, 4, 6, 8, 10, 12\}$ are reduced sets of residues modulo 7. The sets $\{1, 3, 5, 7\}$ and $\{3, 9, 15, 21\}$ are reduced sets of residues modulo 8.

An integer a is called *invertible modulo m* or a *unit modulo m* if there exists an integer x such that

$$ax \equiv 1 \pmod{m}.$$

By Theorem 2.2, a is invertible modulo m if and only if a is relatively prime to m. Moreover, if a is invertible and $ax \equiv 1 \pmod{m}$, then x is unique modulo m. The congruence class $a + m\mathbf{Z}$ is called *invertible* if there exists a congruence class $x + m\mathbf{Z}$ such that

$$(a + m\mathbf{Z})(x + m\mathbf{Z}) = 1 + m\mathbf{Z}.$$

We denote the inverse of the congruence class $a + m\mathbf{Z}$ by $(a + m\mathbf{Z})^{-1} = a^{-1} + m\mathbf{Z}$. The invertible congruence classes are the units in the ring $\mathbf{Z}/m\mathbf{Z}$. We denote the group of units in $\mathbf{Z}/m\mathbf{Z}$ by

$$(\mathbf{Z}/m\mathbf{Z})^{\times}.$$

If $R = \{r_1, \ldots, r_{\varphi(m)}\}$ is a reduced set of residues modulo m, then

$$(\mathbf{Z}/m\mathbf{Z})^{\times} = \{r + m\mathbf{Z} : r \in R\}$$

and

$$\left|(\mathbf{Z}/m\mathbf{Z})^{\times}\right| = \varphi(m).$$

For example,

$$(\mathbf{Z}/6\mathbf{Z})^{\times} = \{1 + 6\mathbf{Z}, 5 + 6\mathbf{Z}\}$$

and

$$(\mathbf{Z}/7\mathbf{Z})^{\times} = \{1 + 7\mathbf{Z}, 2 + 7\mathbf{Z}, 3 + 7\mathbf{Z}, 4 + 7\mathbf{Z}, 5 + 7\mathbf{Z}, 6 + 7\mathbf{Z}\}.$$

If $a + m\mathbf{Z}$ is a unit in $\mathbf{Z}/m\mathbf{Z}$, then $(a, m) = 1$ and we can apply the Euclidean algorithm to compute $(a + m\mathbf{Z})^{-1}$. If we can find integers x and y such that

$$ax + my = 1,$$

then

$$(a + m\mathbf{Z})(x + m\mathbf{Z}) = 1 + m\mathbf{Z},$$

and $x + m\mathbf{Z} = (a + m\mathbf{Z})^{-1}$.

For example, to find the inverse of $13 + 17\mathbf{Z}$, we use the Euclidean algorithm to obtain

$$\begin{aligned} 17 &= 13 \cdot 1 + 4, \\ 13 &= 4 \cdot 3 + 1, \\ 4 &= 1 \cdot 4. \end{aligned}$$

This gives

$$1 = 13 - 4 \cdot 3 = 13 - (17 - 13 \cdot 1)3 = 13 \cdot 4 - 17 \cdot 3,$$

and so

$$13 \cdot 4 \equiv 1 \pmod{17}.$$

Therefore,

$$(13 + 17\mathbf{Z})^{-1} = 4 + 17\mathbf{Z}.$$

Exercises

1. Find all solutions of the congruence $4x \equiv 9 \pmod{11}$.

2. Find all solutions of the congruence $12x \equiv 3 \pmod{45}$.

3. Find all solutions of the congruence $28x \equiv 35 \pmod{42}$.

4. Find all solutions of the system of congruences

$$5x + 7y \equiv 3 \pmod{17}$$

$$2x + 3y \equiv -2 \pmod{17}.$$

5. Find all solutions of the system of congruences

$$8x + 5y \equiv 1 \pmod{13}$$

$$4x + 3y \equiv 3 \pmod{13}.$$

6. Find the inverse of each nonzero congruence class modulo 13.

7. Prove that if m is composite and $m \neq 4$, then $(m-1)! \equiv 0 \pmod{m}$. This is the converse of Wilson's theorem.

8. Prove that if $p \geq 5$ is an odd prime, then

$$6(p-4)! \equiv 1 \pmod{p}.$$

9. Let m and a be integers such that $m \geq 1$ and $(a, m) = 1$. Prove that if $\{r_1, \ldots, r_{\varphi(m)}\}$ is a reduced set of residues modulo m, then $\{ar_1, \ldots, ar_{\varphi(m)}\}$ is also a reduced set of residues modulo m.

10. We say that an integer a is nilpotent modulo m if there exists a positive integer k such that $a^k \equiv 0 \pmod{m}$. Prove that a is nilpotent modulo m if and only if $a \equiv 0 \pmod{\mathrm{rad}(m)}$.

11. For $n \geq 1$, consider the rational number

$$h_n = \sum_{k=1}^{n} \frac{1}{k} = \frac{u_n}{v_n},$$

where u_n and v_n are positive integers. Prove that if p is an odd prime, then the numerator u_{p-1} of h_{p-1} is divisible by p.

Hint: Write h_{p-1} as a fraction with denominator $(p-1)!$, and apply Wilson's theorem.

12. (A criterion for divisibility by 7.) Let n be a positive integer, and let $d_k d_{k-1} \ldots d_1 d_0$ be the usual 10-adic representation of n. Define $f(n) = d_k d_{k-1} \ldots d_1 - 2d_0$. (For example, if $n = 203$, then $d_0 = 3$, $d_1 = 0$, $d_2 = 2$, and $f(203) = 20-6 = 14$.) Prove that n is divisible by 7 if and only if $f(n)$ is divisible by 7. Use this criterion to determine if 7875 is divisible by 7.

    Hint: Prove that $10v + u \equiv 0 \pmod 7$ if and only if $v - 2u \equiv 0 \pmod 7$.

13. Let $k \geq 3$. Find all solutions of the congruence

$$x^2 \equiv 1 \pmod{2^k}.$$

2.3 The Euler Phi Function

An *arithmetic function* is a function defined on the positive integers. The Euler phi function $\varphi(m)$ is the arithmetic function that counts the number of integers in the set $0, 1, 2, \ldots, m-1$ that are relatively prime to m. We have

$$\begin{array}{rclcrcl} \varphi(1) & = & 1, & \quad & \varphi(6) & = & 2, \\ \varphi(2) & = & 2, & & \varphi(7) & = & 6, \\ \varphi(3) & = & 3, & & \varphi(8) & = & 4, \\ \varphi(4) & = & 2, & & \varphi(9) & = & 6, \\ \varphi(5) & = & 4, & & \varphi(10) & = & 4. \end{array}$$

If p is a prime number, then $(a, p) = 1$ for $a = 1, \ldots, p-1$, and $\varphi(p) = p-1$. If p^r is a prime power and $0 \leq a \leq p^r - 1$, then $(a, p^r) > 1$ if and only if a is a multiple of p. The integral multiples of p in the interval $[0, p^r - 1]$ are the p^{r-1} numbers $0, p, 2p, 3p, \ldots, (p^{r-1} - 1)p$, and so

$$\varphi(p^r) = p^r - p^{r-1} = p^r \left(1 - \frac{1}{p}\right).$$

In this section we shall obtain some important properties of the Euler phi function.

Theorem 2.6 *Let m and n be relatively prime positive integers. For every integer c there exist unique integers a and b such that*

$$0 \leq a \leq n-1,$$

$$0 \leq b \leq m-1,$$

and

$$c \equiv ma + nb \pmod{mn}. \tag{2.4}$$

Moreover, $(c, mn) = 1$ if and only if $(a, n) = (b, m) = 1$ in the representation (2.4).

Proof. If a_1, a_2, b_1, b_2 are integers such that

$$ma_1 + nb_1 \equiv ma_2 + nb_2 \pmod{mn},$$

then

$$ma_1 \equiv ma_1 + nb_1 \equiv ma_2 + nb_2 \equiv ma_2 \pmod{n}.$$

Since $(m, n) = 1$, it follows that

$$a_1 \equiv a_2 \pmod{n},$$

and so $a_1 = a_2$. Similarly, $b_1 = b_2$. It follows that the mn integers $ma + nb$ are pairwise incongruent modulo mn. Since there are exactly mn distinct congruence classes modulo mn, the congruence (2.4) has a unique solution for every integer c.

Let $c \equiv ma + nb \pmod{mn}$. Since $(m, n) = 1$, we have

$$(c, m) = (ma + nb, m) = (nb, m) = (b, m)$$

and

$$(c, n) = (ma + nb, n) = (ma, n) = (a, n).$$

It follows that $(c, mn) = 1$ if and only if $(c, m) = (c, n) = 1$ if and only if $(b, m) = (a, n) = 1$. This completes the proof. □

For example, we can represent the congruence classes modulo 6 as linear combinations of 2 and 3 as follows:

$$\begin{aligned}
0 &\equiv 0 \cdot 2 + 0 \cdot 3 \pmod{6}, \\
1 &\equiv 2 \cdot 2 + 1 \cdot 3 \pmod{6}, \\
2 &\equiv 1 \cdot 2 + 0 \cdot 3 \pmod{6}, \\
3 &\equiv 0 \cdot 2 + 1 \cdot 3 \pmod{6}, \\
4 &\equiv 2 \cdot 2 + 0 \cdot 3 \pmod{6}, \\
5 &\equiv 1 \cdot 2 + 1 \cdot 3 \pmod{6}.
\end{aligned}$$

A *multiplicative* function is an arithmetic function $f(m)$ such that $f(mn) = f(m)f(n)$ for all pairs of relatively prime positive integers m and n. If $f(m)$ is multiplicative, then it is easy to prove by induction on k that if $m_1, \ldots, m_k$ are pairwise relatively prime positive integers, then $f(m_1 \cdots m_k) = f(m_1) \cdots f(m_k)$.

Theorem 2.7 *The Euler phi function is multiplicative. Moreover,*

$$\varphi(m) = m \prod_{p|m} \left(1 - \frac{1}{p}\right).$$

Proof. Let $(m,n)=1$. There are $\varphi(mn)$ congruence classes in the ring $\mathbf{Z}/mn\mathbf{Z}$ that are relatively prime to mn. By Theorem 2.6, every congruence class modulo mn can be written uniquely in the form $ma+nb+mn\mathbf{Z}$, where a and b are integers such that $0 \leq a \leq n-1$ and $0 \leq b \leq m-1$. Moreover, the congruence class $ma+nb+mn\mathbf{Z}$ is prime to mn if and only if $(b,m)=(a,n)=1$. Since there are $\varphi(n)$ integers $a \in [0,n-1]$ that are relatively prime to n, and $\varphi(m)$ integers $b \in [0,m-1]$ relatively prime to m, it follows that $\varphi(mn)=\varphi(m)\varphi(n)$, and so the Euler phi function is multiplicative. If $m_1,\ldots,m_k$ are pairwise relatively prime positive integers, then $\varphi(m_1\cdots m_k)=\varphi(m_1)\cdots\varphi(m_k)$. In particular, if $m=p_1^{r_1}\cdots p_k^{r_k}$ is the standard factorization of m, where $p_1,\ldots,p_k$ are distinct primes and $r_1,\ldots,r_k$ are positive integers, then

$$\varphi(m)=\prod_{i=1}^{k}\varphi\left(p_i^{r_i}\right)=\prod_{i=1}^{k}p_i^{r_i}\left(1-\frac{1}{p_i}\right)=m\prod_{p|m}\left(1-\frac{1}{p}\right).$$

This completes the proof. □

For example, $7875=3^2 5^3 7$ and

$$\varphi(7875)=\varphi(3^2)\varphi(5^3)\varphi(7)=(9-3)(125-25)(7-1)=3600.$$

Theorem 2.8 *For every positive integer* m,

$$\sum_{d|m}\varphi(d)=m.$$

Proof. We first consider the case where $m=p^t$ is a power of a prime p. The divisors of p^t are $1,p,p^2,\ldots,p^t$, and

$$\sum_{d|p^t}\varphi(d)=\sum_{r=0}^{t}\varphi(p^r)=1+\sum_{r=1}^{t}\left(p^r-p^{r-1}\right)=p^t.$$

Next we consider the general case where m has the standard factorization

$$m=p_1^{t_1}p_2^{t_2}\cdots p_k^{t_k},$$

where $p_1,\ldots,p_k$ are distinct prime numbers and $t_1,\ldots,t_k$ are positive integers. Every divisor d of m is of the form

$$d=p_1^{r_1}p_2^{r_2}\cdots p_k^{r_k},$$

where $0 \leq r_i \leq t_i$ for $i=1,\ldots,k$. By Theorem 2.7, $\varphi(d)$ is multiplicative, and so

$$\varphi(d)=\varphi(p_1^{r_1})\varphi(p_2^{r_2})\cdots\varphi(p_k^{r_k}).$$

Therefore,

$$\begin{aligned}
\sum_{d|m} \varphi(d) &= \sum_{r_1=0}^{t_1} \cdots \sum_{r_k=0}^{t_k} \varphi\left(p_1^{r_1} \cdots p_k^{r_k}\right) \\
&= \sum_{r_1=0}^{t_1} \cdots \sum_{r_k=0}^{t_k} \varphi(p_1^{r_1})\varphi(p_2^{r_2}) \cdots \varphi(p_k^{r_k}) \\
&= \prod_{i=1}^{k} \sum_{r_i=0}^{t_i} \varphi\left(p_i^{r_i}\right) \\
&= \prod_{i=1}^{k} p_i^{t_i} \\
&= m.
\end{aligned}$$

This completes the proof. □

For example,

$$\begin{aligned}
\sum_{d|12} \varphi(d) &= \varphi(1) + \varphi(2) + \varphi(3) + \varphi(4) + \varphi(6) + \varphi(12) \\
&= 1 + 1 + 2 + 2 + 2 + 4 \\
&= 12
\end{aligned}$$

and

$$\begin{aligned}
\sum_{d|45} \varphi(d) &= \varphi(1) + \varphi(3) + \varphi(5) + \varphi(9) + \varphi(15) + \varphi(45) \\
&= 1 + 2 + 4 + 6 + 8 + 24 \\
&= 45.
\end{aligned}$$

Exercises

1. Compute $\varphi(6993)$.

2. Represent the congruence classes modulo 12 in the form $3a + 4b$ with $0 \leq a \leq 3$ and $0 \leq b \leq 2$.

3. Let $m = 15$. Compute $\varphi(d)$ for every divisor d of m, and check that $\sum_{d|m} \varphi(d) = m$. Repeat this exercise for $m = 16, 17$, and 18.

4. Prove that $\varphi(m)$ is even for all $m \geq 3$.

5. Prove that $\varphi(m^k) = m^{k-1}\varphi(m)$ for all positive integers m and k.

6. Prove that m is prime if and only if $\varphi(m) = m - 1$.

7. Prove that $\varphi(m) = \varphi(2m)$ if and only if m is odd.

8. Prove that if m divides n, then $\varphi(m)$ divides $\varphi(n)$.

9. Find all positive integers n such that $\varphi(n)$ is not divisible by 4.

10. Find all positive integers n such that $\varphi(5n) = 5\varphi(n)$.

11. Let $f(n) = \varphi(n)/n$. Prove that $\varphi(p^k) = \varphi(p)$ for all primes p and all positive integers k.

12. This problem gives an alternative proof of Theorem 2.8. Let $m \geq 1$, and let S be the set of fractions k/m with $k = 0, 1, \ldots, m-1$. Write each fraction in lowest terms: $k/m = a/d$, where d is a divisor of m and $(a, d) = 1$. For example, $0/m = 0/1$. Show that for each divisor d of m there are exactly $\varphi(d)$ fractions $k/m \in S$ that have denominator d when reduced to lowest terms. Deduce that $\sum_{d|m} \varphi(d) = m$.

13. Let $N_m(x)$ denote the number of positive integers not exceeding x that are relatively prime to m. Prove that

$$\lim_{x \to \infty} \frac{N_m(x)}{x} = \frac{\varphi(m)}{m}.$$

This result can be expressed as follows: The probability that a random integer is prime to m is $\varphi(m)/m$.

2.4 Chinese Remainder Theorem

Theorem 2.9 *Let m and n be positive integers. For any integers a and b there exists an integer x such that*

$$x \equiv a \pmod{m} \tag{2.5}$$

and

$$x \equiv b \pmod{n} \tag{2.6}$$

if and only if

$$a \equiv b \pmod{(m, n)}.$$

If x is a solution of congruences (2.5) and (2.6), then the integer y is also a solution if and only if

$$x \equiv y \pmod{[m, n]}.$$

Proof. If x is a solution of congruence (2.5), then $x = a + mu$ for some integer u. If x is also a solution of congruence (2.6), then

$$x = a + mu \equiv b \pmod{n},$$

that is,

$$a + mu = b + nv$$

for some integer v. It follows that

$$a - b = nv - mu \equiv 0 \pmod{(m, n)}.$$

Conversely, if $a - b \equiv 0 \pmod{(m, n)}$, then by Theorem 1.15 there exist integers u and v such that

$$a - b = nv - mu.$$

Then

$$x = a + mu = b + nv$$

is a solution of the two congruences.

An integer y is another solution of the congruences if and only if

$$y \equiv a \equiv x \pmod{m}$$

and

$$y \equiv b \equiv x \pmod{n},$$

that is, if and only if $x-y$ is a common multiple of m and n, or, equivalently, $x - y$ is divisible by the least common multiple $[m, n]$. This completes the proof. □

For example, the system of congruences

$$\begin{aligned} x &\equiv 5 \pmod{21}, \\ x &\equiv 19 \pmod{56}, \end{aligned}$$

has a solution, since

$$(56, 21) = 7$$

and

$$19 \equiv 5 \pmod{7}.$$

The integer x is a solution if there exists an integer u such that

$$x = 5 + 21u \equiv 19 \pmod{56},$$

that is,

$$21u \equiv 14 \pmod{56},$$

$$3u \equiv 2 \pmod{8},$$

or

$$u \equiv 6 \pmod{8}.$$

Then

$$x = 5 + 21u = 5 + 21(6 + 8v) = 131 + 168v$$

is a solution of the system of congruences for any integer v, and so the set of all solutions is the congruence class $131 + 168\mathbf{Z}$.

Theorem 2.10 (Chinese remainder theorem) *Let $k \geq 2$. If $a_1, \ldots, a_k$ are integers and $m_1, \ldots, m_k$ are pairwise relatively prime positive integers, then there exists an integer x such that*

$$x \equiv a_i \pmod{m_i} \quad \text{for all } i = 1, \ldots, k.$$

If x is any solution of this set of congruences, then the integer y is also a solution if and only if

$$x \equiv y \pmod{m_1 \cdots m_k}.$$

Proof. We prove the theorem by induction on k. If $k = 2$, then $[m_1, m_2] = m_1 m_2$, and this is a special case of Theorem 2.9.

Let $k \geq 3$, and assume that the statement is true for $k - 1$ congruences. Then there exists an integer z such that $z \equiv a_i \pmod{m_i}$ for $i = 1, \ldots, k-1$. Since $m_1, \ldots, m_k$ are pairwise relatively prime integers, we have

$$(m_1 \cdots m_{k-1}, m_k) = 1,$$

and so, by the case $k = 2$, there exists an integer x such that

$$\begin{aligned} x &\equiv z \pmod{m_1 \cdots m_{k-1}}, \\ x &\equiv a_k \pmod{m_k}. \end{aligned}$$

Then

$$x \equiv z \equiv a_i \pmod{m_i}$$

for $i = 1, \ldots, k - 1$.

If y is another solution of the system of k congruences, then $x - y$ is divisible by m_i for all $i = 1, \ldots, k$. Since $m_1, \ldots, m_k$ are pairwise relatively prime, it follows that $x - y$ is divisible by $m_1 \cdots m_k$. This completes the proof. □

For example, the system of congruences

$$\begin{aligned} x &\equiv 2 \pmod{3}, \\ x &\equiv 3 \pmod{5}, \\ x &\equiv 5 \pmod{7}, \\ x &\equiv 7 \pmod{11} \end{aligned}$$

has a solution, since the moduli are pairwise relatively prime. The solution to the first two congruences is the congruence class

$$x \equiv 8 \pmod{15}.$$

The solution to the first three congruences is the congruence class

$$x \equiv 68 \pmod{105}.$$

The solution to the four congruences is the congruence class

$$x \equiv 1118 \pmod{1155}.$$

There is an important application of the Chinese remainder theorem to the problem of solving diophantine equations of the form

$$f(x_1, \ldots, x_k) \equiv 0 \pmod{m},$$

where $f(x_1, \ldots, x_k)$ is a polynomial with integer coefficients in one or several variables. This equation is *solvable modulo m* if there exist integers $a_1, \ldots, a_k$ such that

$$f(a_1, \ldots, a_k) \equiv 0 \pmod{m}.$$

The Chinese remainder theorem allows us to reduce the question of the solvability of this congruence modulo m to the special case of prime power moduli p^r. For simplicity, we consider polynomials in only one variable.

Theorem 2.11 *Let*

$$m = p_1^{r_1} \cdots p_k^{r_k}$$

be the standard factorization of the positive integer m. Let $f(x)$ be a polynomial with integral coefficients. The congruence

$$f(x) \equiv 0 \pmod{m}$$

is solvable if and only if the congruences

$$f(x) \equiv 0 \pmod{p_i^{r_i}}$$

are solvable for all $i = 1, \ldots, k$.

Proof. If $f(x) \equiv 0 \pmod{m}$ has a solution in integers, then there exists an integer a such that m divides $f(a)$. Since $p_i^{r_i}$ divides m, it follows that $p_i^{r_i}$ divides $f(a)$, and so the congruences $f(x) \equiv 0 \pmod{p_i^{r_i}}$ are solvable for $i = 1, \ldots, k$.

Conversely, suppose that the congruences $f(x) \equiv 0 \pmod{p_i^{r_i}}$ are solvable for $i = 1, \ldots, k$. Then for each i there exists an integer a_i such that

$$f(a_i) \equiv 0 \pmod{p_i^{r_i}}.$$

Since the prime powers $p_1^{r_1}, \ldots, p_k^{r_k}$ are pairwise relatively prime, the Chinese remainder theorem tells us that there exists an integer a such that

$$a \equiv a_i \pmod{p_i^{r_i}}$$

for all i. Then

$$f(a) \equiv f(a_i) \equiv 0 \pmod{p_i^{r_i}}$$

for all i. Since $f(a)$ is divisible by each of the prime powers $p_i^{r_i}$, it is also divisible by their product m, and so $f(a) \equiv 0 \pmod{m}$. This completes the proof. □

For example, consider the congruence

$$f(x) = x^2 - 34 \equiv 0 \pmod{495}.$$

Since $495 = 3^2 \cdot 5 \cdot 11$, it suffices to solve the congruences

$$f(x) = x^2 - 34 \equiv x^2 + 2 \equiv 0 \pmod{9},$$

$$f(x) = x^2 - 34 \equiv x^2 + 1 \equiv 0 \pmod{5},$$

and

$$f(x) = x^2 - 34 \equiv x^2 - 1 \equiv 0 \pmod{11}.$$

These congruences have solutions

$$f(5) \equiv 0 \pmod{9},$$

$$f(2) \equiv 0 \pmod{5},$$

and

$$f(1) \equiv 0 \pmod{11}.$$

By the Chinese remainder theorem, there exists an integer a such that

$$\begin{aligned} a &\equiv 5 \pmod{9}, \\ a &\equiv 2 \pmod{5}, \\ a &\equiv 1 \pmod{11}. \end{aligned}$$

Solving these congruences, we obtain

$$a \equiv 122 \pmod{495}.$$

We can check that

$$f(122) = 122^2 - 34 = 14,850 = 30 \cdot 495,$$

and so

$$f(122) \equiv 0 \pmod{495}.$$

Exercises

1. Find all solutions of the system of congruences

$$\begin{aligned} x &\equiv 4 \pmod{5}, \\ x &\equiv 5 \pmod{6}. \end{aligned}$$

2. Find all solutions of the system of congruences

$$\begin{aligned} x &\equiv 5 \pmod{12}, \\ x &\equiv 8 \pmod{9}. \end{aligned}$$

3. Find all solutions of the system of congruences

$$\begin{aligned} x &\equiv 5 \pmod{12}, \\ x &\equiv 8 \pmod{10}. \end{aligned}$$

4. Find all solutions of the system of congruences

$$\begin{aligned} 2x &\equiv 1 \pmod{5}, \\ 3x &\equiv 4 \pmod{7}. \end{aligned}$$

5. Find all integers that have a remainder of 1 when divided by 3, 5, and 7.

6. Find all integers that have a remainder of 2 when divided by 4 and that have a remainder of 3 when divided by 5.

7. Find all solutions of the congruence

$$f(x) = 5x^3 - 93 \equiv 0 \pmod{231}.$$

8. (Bhaskara, sixth century) A basket contains n eggs. If the eggs are removed $2, 3, 4, 5$, or 6 at a time, then the number of eggs that remain in the basket is $1, 2, 3, 4$, or 5, respectively. If the eggs are removed 7 at a time, then no eggs remain. What is the smallest number n of eggs that could have been in the basket at the start of this procedure?

   Hint: The first condition implies that $n \equiv 1 \pmod{2}$.

9. Let f be a polynomial with integer coefficients. For $m \geq 1$, let $N_f(m)$ denote the number of pairwise incongruent solutions of $f(x) \equiv 0 \pmod{m}$. Prove that the function $N_f(m)$ is multiplicative, that is, $N_f(m_1m_2) = N_f(m_1)N_f(m_2)$ if $(m_1, m_2) = 1$.

10. Let $m_1, \ldots, m_k$ be pairwise relatively prime positive integers and $m = m_1 \cdots m_k$. Define the map
$$f : (\mathbf{Z}/m\mathbf{Z})^\times \to (\mathbf{Z}/m_1\mathbf{Z})^\times \times \cdots \times (\mathbf{Z}/m_k\mathbf{Z})^\times$$
by
$$f(a + m\mathbf{Z}) = (a + m_1\mathbf{Z}, \ldots, a + m_k\mathbf{Z}).$$
Use the Chinese remainder theorem to show directly that this map is one-to-one and onto.

2.5 Euler's Theorem and Fermat's Theorem

Euler's theorem and its corollary, Fermat's theorem, are fundamental results in number theory, with many applications in mathematics and computer science. In the following sections we shall see how the Euler and Fermat theorems can be used to determine whether an integer is prime or composite, and how they are applied in cryptography.

Theorem 2.12 (Euler) *Let m be a positive integer, and let a be an integer relatively prime to m. Then*
$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

Proof. Let $\{r_1, \ldots, r_{\varphi(m)}\}$ be a reduced set of residues modulo m. Since $(a, m) = 1$, we have $(ar_i, m) = 1$ for $i = 1, \ldots, \varphi(m)$. Consequently, for every $i \in \{1, \ldots, \varphi(m)\}$ there exists $\sigma(i) \in \{1, \ldots, \varphi(m)\}$ such that
$$ar_i \equiv r_{\sigma(i)} \pmod{m}.$$
Moreover, $ar_i \equiv ar_j \pmod{m}$ if and only if $i = j$, and so σ is a permutation of the set $\{1, \ldots, \varphi(m)\}$ and $\{ar_1, \ldots, ar_{\varphi(m)}\}$ is also a reduced set of residues modulo m. It follows that
$$\begin{aligned} a^{\varphi(m)} r_1 r_2 \cdots r_{\varphi(m)} &\equiv (ar_1)(ar_2) \cdots (ar_{\varphi(m)}) \pmod{m} \\ &\equiv r_{\sigma(1)} r_{\sigma(2)} \cdots r_{\sigma(\varphi(m))} \pmod{m} \\ &\equiv r_1 r_2 \cdots r_{\varphi(m)} \pmod{m}. \end{aligned}$$
Dividing by $r_1 r_2 \cdots r_{\varphi(m)}$, we obtain
$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$
This completes the proof. □

The following corollary is sometimes called *Fermat's little theorem.*

Theorem 2.13 (Fermat) *Let p be a prime number. If the integer a is not divisible by p, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Moreover,

$$a^p \equiv a \pmod{p}$$

for every integer a.

Proof. If p is prime and does not divide a, then $(a,p)=1$, $\varphi(p)=p-1$, and

$$a^{p-1} = a^{\varphi(p)} \equiv 1 \pmod{p}$$

by Euler's theorem. Multiplying this congruence by a, we obtain

$$a^p \equiv a \pmod{p}.$$

If p divides a, then this congruence also holds for a. □

Let m be a positive integer and let a be an integer that is relatively prime to m. By Euler's theorem, $a^{\varphi(m)} \equiv 1 \pmod{m}$. The *order* of a with respect to the modulus m is the smallest positive integer d such that $a^d \equiv 1 \pmod{m}$. Then $1 \leq d \leq \varphi(m)$. We denote the order of a modulo m by $\mathrm{ord}_m(a)$. We shall prove that $\mathrm{ord}_m(a)$ divides $\varphi(m)$ for every integer a relatively prime to p.

Theorem 2.14 *Let m be a positive integer and a an integer relatively prime to m. If d is the order of a modulo m, then $a^k \equiv a^\ell \pmod{m}$ if and only if $k \equiv \ell \pmod{d}$. In particular, $a^n \equiv 1 \pmod{m}$ if and only if d divides n, and so d divides $\varphi(m)$.*

Proof. Since a has order d modulo m, we have $a^d \equiv 1 \pmod{m}$. If $k \equiv \ell \pmod{d}$, then $k = \ell + dq$, and so

$$a^k = a^{\ell+dq} = a^\ell \left(a^d\right)^q \equiv a^\ell \pmod{m}.$$

Conversely, suppose that $a^k \equiv a^\ell \pmod{m}$. By the division algorithm, there exist integers q and r such that

$$k - \ell = dq + r \qquad \text{and} \qquad 0 \leq r \leq d-1.$$

Then

$$a^k = a^{\ell+dq+r} = a^\ell \left(a^d\right)^q a^r \equiv a^k a^r \pmod{m}.$$

Since $(a^k, m) = 1$, we can divide this congruence by a^k and obtain

$$a^r \equiv 1 \pmod{m}.$$

Since $0 \leq r \leq d-1$, and d is the order of a modulo m, it follows that $r = 0$, and so $k \equiv \ell \pmod{d}$.

If $a^n \equiv 1 \equiv a^0 \pmod{m}$, then d divides n. In particular, d divides $\varphi(m)$, since $a^{\varphi(m)} \equiv 1 \pmod{m}$ by Euler's theorem. □

For example, let $m = 15$ and $a = 7$. Since $\varphi(15) = 8$, Euler's theorem tells us that

$$7^8 \equiv 1 \pmod{15}.$$

Moreover, the order of 7 with respect to 15 is a divisor of 8. We can compute the order as follows:

$$\begin{aligned} 7^1 &\equiv 7 \pmod{15}, \\ 7^2 &\equiv 49 \equiv 4 \pmod{15}, \\ 7^3 &\equiv 28 \equiv 13 \pmod{15}, \\ 7^4 &\equiv 91 \equiv 1 \pmod{15}, \end{aligned}$$

and so the order of 7 is 4.

We shall give a second proof of Euler's theorem and its corollaries. We begin with some simple observations about groups. We define the *order of a group* as the cardinality of the group.

Theorem 2.15 (Lagrange's theorem) *If G is a finite group and H is a subgroup of G, then the order of H divides the order of G.*

Proof. Let G be a group, written multiplicatively, and let X be a nonempty subset of G. For every $a \in G$ we define the set

$$aX = \{ax : x \in X\}.$$

The map $f : X \to aX$ defined by $f(x) = ax$ is a bijection, and so $|X| = |aX|$ for all $a \in G$. If H is a subgroup of G, then aH is called a *coset* of H. Let aH and bH be cosets of the subgroup H. If $aH \cap bH \neq \emptyset$, then there exist $x, y \in H$ such that $ax = by$, or, since H is a subgroup, $b = axy^{-1} = az$, where $z = xy^{-1} \in H$. Then $bh = azh \in aH$ for all $h \in H$, and so $bH \subseteq aH$. By symmetry, $aH \subseteq bH$, and so $aH = bH$. Therefore, cosets of a subgroup H are either disjoint or equal. Since every element of G belongs to some coset of H (for example, $a \in aH$ for all $a \in G$), it follows that the cosets of H partition G. We denote the set of cosets by G/H. If G is a finite group, then H and G/H are finite, and

$$|G| = |H||G/H|.$$

In particular, we see that $|H|$ divides $|G|$. □

Let G be a group, written multiplicatively, and let $a \in G$. Let $H = \{a^k : k \in \mathbf{Z}\}$. Then $1 = a^0 \in H \subseteq G$. Since $a^k a^\ell = a^{k+\ell}$ for all $k, \ell \in \mathbf{Z}$, it follows

that H is a subgroup of G. This subgroup is called the *cyclic subgroup generated by* a, and written $\langle a\rangle$. Cyclic subgroups are abelian.

The group G is *cyclic* if there exists an element $a \in G$ such that $G = \langle a\rangle$. In this case, the element a is called a *generator* of G. For example, the group $(\mathbf{Z}/7\mathbf{Z})^{\times}$ is a cyclic group of order 6 generated by $3 + 7\mathbf{Z}$. The congruence class $5 + 7\mathbf{Z}$ is another generator of this group.

If $a^k \neq a^\ell$ for all integers $k \neq \ell$, then the cyclic subgroup generated by a is infinite. If there exist integers k and ℓ such that $k < \ell$ and $a^k = a^\ell$, then $a^{\ell-k} = 1$. Let d be the smallest positive integer such that $a^d = 1$. Then the group elements $1, a, a^2, \ldots, a^{d-1}$ are distinct. Let $n \in \mathbf{Z}$. By the division algorithm, there exist integers q and r such that $n = dq + r$ and $0 \leq r \leq d-1$. Since

$$a^n = a^{dq+r} = \left(a^d\right)^q a^r = a^r,$$

it follows that

$$\langle a\rangle = \{a^n : n \in \mathbf{Z}\} = \{a^r : 0 \leq r \leq d-1\},$$

and the cyclic subgroup generated by a has order d. Moreover, $a^k = a^\ell$ if and only if $k \equiv \ell \pmod{d}$.

Let G be a group, and let $a \in G$. We define the *order* of a as the cardinality of the cyclic subgroup generated by a.

Theorem 2.16 *Let G be a finite group, and $a \in G$. Then the order of the element a divides the order of the group G.*

Proof. This follows immediately from Theorem 2.15, since the order of a is the order of the cyclic subgroup that a generates. □

Let us apply these remarks to the special case when $G = (\mathbf{Z}/m\mathbf{Z})^{\times}$ is the group of units in the ring of congruence classes modulo m. Then G is a finite group of order $\varphi(m)$. Let $(a, m) = 1$ and let d be the order of $a + m\mathbf{Z}$ in G, that is, the order of the cyclic subgroup generated by $a + m\mathbf{Z}$. By Theorem 2.16, d divides $\varphi(m)$, and so

$$a^{\varphi(m)} + m\mathbf{Z} = (a + m\mathbf{Z})^{\varphi(m)} = \left((a + m\mathbf{Z})^d\right)^{\varphi(m)/d} = 1 + m\mathbf{Z}.$$

Equivalently,

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

This is Euler's theorem.

Theorem 2.17 *Let G be a cyclic group of order m, and let H be a subgroup of G. If a is a generator of G, then there exists a unique divisor d of m such that H is the cyclic subgroup generated by a^d, and H has order m/d.*

Proof. Let S be the set of all integers u such that $a^u \in H$. If $u, v \in S$, then $a^u, a^v \in H$. Since H is a subgroup, it follows that $a^u a^v = a^{u+v} \in H$ and $a^u(a^v)^{-1} = a^{u-v} \in H$. Therefore, $u \pm v \in S$, and S is a subgroup of $\mathbf{Z}$. By Theorem 1.3, there is a unique nonnegative integer d such that $S = d\mathbf{Z}$, and so H is the cyclic subgroup generated by a^d. Since $a^m = 1 \in H$, we have $m \in S$, and so d is a positive divisor of m. It follows that H has order m/d. □

Theorem 2.18 *Let G be a cyclic group of order m, and let a be a generator of G. For every integer k, the cyclic subgroup generated by a^k has order m/d, where $d = (m, k)$, and $\langle a^k \rangle = \langle a^d \rangle$. In particular, G has exactly $\varphi(m)$ generators.*

Proof. Since $d = (k, m)$, there exist integers x and y such that $d = kx + my$. Then

$$a^d = a^{kx+my} = \left(a^k\right)^x (a^m)^y = \left(a^k\right)^x,$$

and so $a^d \in \langle a^k \rangle$ and $\langle a^d \rangle \subseteq \langle a^k \rangle$. Since d divides k, there exists an integer z such that $k = dz$. Then

$$a^k = \left(a^d\right)^z,$$

and so $a^k \in \langle a^d \rangle$ and $\langle a^k \rangle \subseteq \langle a^d \rangle$. Therefore, $\langle a^k \rangle = \langle a^d \rangle$ and a^k has order m/d. In particular, a^k generates G if and only if $d = 1$ if and only if $(m, k) = 1$, and so G has exactly $\varphi(m)$ generators. This completes the proof. □

We can now give a group theoretic proof of Theorem 2.8. Let G be a cyclic group of order m. For every divisor d of m, the group G has a unique cyclic subgroup of order d, and this subgroup has exactly $\varphi(d)$ generators. Since every element of G generates a cyclic subgroup, it follows that

$$m = \sum_{d|m} \varphi(d).$$

Voilà!

Exercises

1. Prove that
$$3^{512} \equiv 1 \pmod{1024}.$$

2. Find the remainder when 7^{51} is divided by 144.

3. Find the remainder when 2^{10^8} is divided by 31.

4. Compute the order of 2 with respect to the prime moduli 3, 5, 7, 11, 13, 17, and 19.

5. Compute the order of 10 with respect to the modulus 7.

6. Let r_i denote the least nonnegative residue of $10^i \pmod 7$. Compute r_i for $i = 1, \ldots, 6$. Compute the decimal expansion of the fraction $1/7$ without using a calculator. Can you find where the numbers $r_1, \ldots, r_6$ appear in the process of dividing 7 into 1?

7. Compute the order of 10 modulo 13. Compute the period of the fraction $1/13$.

8. Let p be prime and a an integer not divisible by p. Prove that if $a^{2^n} \equiv -1 \pmod p$, then a has order 2^{n+1} modulo p.

9. Let m be a positive integer not divisible by 2 or 5. Prove that the decimal expansion of the fraction $1/m$ is periodic with period equal to the order of 10 modulo m.

10. Prove that the decimal expansion of $1/m$ is finite if and only if the prime divisors of m are 2 and 5.

11. Prove that 10 has order 22 modulo 23. Deduce that the decimal expansion of $1/23$ has period 22.

12. Prove that if p is a prime number congruent to 1 modulo 4, then there exists an integer x such that $x^2 \equiv -1 \pmod p$.

    Hint: Observe that

$$\begin{aligned}(p-1)! &\equiv \prod_{j=1}^{(p-1)/2} j(p-j) \equiv \prod_{j=1}^{(p-1)/2} (-j^2) \\ &\equiv (-1)^{(p-1)/2} \left(\prod_{j=1}^{(p-1)/2} j \right)^2 \pmod p,\end{aligned}$$

    and apply Theorem 2.4.

13. Prove that if $n \geq 2$, then $2^n - 1$ is not divisible by n.

    Hint: Let p be the smallest prime that divides n. Consider the congruence $2^n \equiv 1 \pmod p$.

14. Prove that if p and q are distinct primes, then

$$p^{q-1} + q^{p-1} \equiv 1 \pmod{pq}.$$

15. Prove that if m and n are relatively prime positive integers, then

$$m^{\varphi(n)} + n^{\varphi(m)} \equiv 1 \pmod{mn}.$$

16. Let p be an odd prime. By Euler's theorem, if $(a, p) = 1$, then

$$f_p(a) = \frac{a^{p-1} - 1}{p} \in \mathbf{Z}.$$

Prove that if $(ab, p) = 1$, then

$$f_p(ab) \equiv f_p(a) + f_p(b) \pmod{p}.$$

17. Let $f(x)$ and $g(x)$ be polynomials with integer coefficients. We say that $f(x)$ is equivalent to $g(x)$ modulo p if

$$f(a) \equiv g(a) \pmod{p} \quad \text{for all integers } a.$$

Prove that the polynomials x^9+5x^7+3 and $x^3-2x+24$ are equivalent modulo 7. Prove that every polynomial is equivalent modulo p to a polynomial of degree at most $p-1$.

Hint: Use Fermat's theorem.

18. Let G be the group $(\mathbf{Z}/7\mathbf{Z})^\times$. Determine all the cyclic subgroups of G.

19. Prove that the group $(\mathbf{Z}/11\mathbf{Z})^\times$ is cyclic, and find a generator.

20. Let G be a group with subgroup H. Define a relation $\sim$ on G as follows: $a \sim b$ if $b^{-1}a \in H$. Prove that this is an equivalence relation (that is, reflexive, symmetric, and transitive). Prove that $a \sim b$ if and only if $aH = bH$, and so the equivalence classes of this relation are the cosets in G/H.

21. Let G be an abelian group with subgroup H. Let G/H be the set of cosets of H in G. Define multiplication of congruence classes by

$$aH \cdot bH = abH.$$

Prove that if $aH = a'H$ and $bH = b'H$, then $abH = a'b'H$, and so multiplication of cosets is well-defined. Prove that G/H is an abelian group with this multiplication. This is called the *quotient group* of G by H.

22. Let G be a group and let H and K be subgroups of G. For $a \in G$, we define the *double coset* $HaK = \{hak : h \in H, k \in K\}$. Prove that if $a, b \in G$ and $HaK \cap HbK \neq \emptyset$, then $HaK = HbK$.

2.6 Pseudoprimes and Carmichael Numbers

Suppose we are given an odd integer $n \geq 3$, and we want to determine whether n is prime or composite. If n is "small," we can simply divide n by all odd integers d such that $3 \leq d \leq \sqrt{n}$. If some d divides n, then n is composite; otherwise, n is prime. If n is "big," however, this method is time-consuming and impractical. We need to find other primality tests.

Fermat's theorem can be applied to this problem. By Fermat's theorem, if n is an odd prime, then $2^{n-1} \equiv 1 \pmod{n}$. Therefore, if n is odd and $2^{n-1} \not\equiv 1 \pmod{n}$, then n must be composite. In general, we can choose any integer b that is relatively prime to n. By Fermat's theorem, if n is prime, then $b^{n-1} \equiv 1 \pmod{n}$. It follows that if $b^{n-1} \not\equiv 1 \pmod{n}$, then n must be composite. Thus, for every base b, Fermat's theorem gives a *primality test*, that is, a necessary condition for an integer n to be prime.

Suppose we want to know whether $n = 851$ is prime or composite. We shall compute $2^{850} \pmod{851}$. An efficient method is to use the 2-adic representation of 850:

$$850 = 2 + 2^4 + 2^6 + 2^8 + 2^9.$$

Since $2^{2^n} = \left(2^{2^{n-1}}\right)^2$, we have

$$\begin{aligned}
2^2 &\equiv 4 \pmod{851}, \\
2^{2^2} &\equiv 16 \pmod{851}, \\
2^{2^3} &\equiv 256 \pmod{851}, \\
2^{2^4} &\equiv 9 \pmod{851}, \\
2^{2^5} &\equiv 81 \pmod{851}, \\
2^{2^6} &\equiv 604 \pmod{851}, \\
2^{2^7} &\equiv 588 \pmod{851}, \\
2^{2^8} &\equiv 238 \pmod{851}, \\
2^{2^9} &\equiv 478 \pmod{851}.
\end{aligned}$$

Then

$$\begin{aligned}
2^{850} &\equiv 2^2 2^{2^4} 2^{2^6} 2^{2^8} 2^{2^9} \pmod{851} \\
&\equiv 4 \cdot 9 \cdot 604 \cdot 238 \cdot 478 \pmod{851} \\
&\equiv 169 \not\equiv 1 \pmod{581},
\end{aligned}$$

and so 851 is composite. To factor 851, we observe that $851 + 49 = 900$, and so

$$851 = 900 - 49 = 30^2 - 7^2 = (30 - 7)(30 + 7) = 23 \cdot 37.$$

(To understand this factoring method, see Exercise 2.)

This test can prove that an integer is composite, but it cannot prove that an integer is prime. For example, consider the composite number $n = 341 = 11 \cdot 31$, Choosing base $b = 2$, we have

$$2^{10} \equiv 1 \pmod{11},$$

and so

$$2^{340} \equiv \left(2^{10}\right)^{34} \equiv 1 \pmod{11}.$$

Similarly,

$$2^{5} \equiv 1 \pmod{31},$$

and so

$$2^{340} \equiv \left(2^{5}\right)^{68} \equiv 1 \pmod{31}.$$

Since $2^{340} - 1$ is divisible by both 11 and 31, it is divisible by their product, that is,

$$2^{340} \equiv 1 \pmod{341}.$$

A composite number n is called a *pseudoprime to the base* b if $(b, n) = 1$ and $b^{n-1} \equiv 1 \pmod{n}$. Thus, 341 is a pseudoprime to base 2.

We can show that 341 is composite by choosing the base $b = 7$. Since

$$7^3 = 343 \equiv 2 \pmod{341}$$

and

$$2^{10} = 1024 \equiv 1 \pmod{341},$$

it follows that

$$\begin{aligned} 7^{340} &= 7\left(7^3\right)^{113} \\ &\equiv 7 \cdot 2^{113} \pmod{341} \\ &\equiv 7 \cdot 2^3 \left(2^{10}\right)^{11} \pmod{341} \\ &\equiv 56 \pmod{341} \\ &\not\equiv 1 \pmod{341}. \end{aligned}$$

Can every composite number be *proved* composite by some primality test based on Fermat's theorem? It is a surprising fact that the answer is "no." There exist composite numbers n that cannot be proved composite by any congruence of the form $b^{n-1} \pmod{n}$ with $(b, n) = 1$. For example, $561 = 3 \cdot 11 \cdot 17$ is composite. Let b be an integer relatively prime to 561. Then

$$b^2 \equiv 1 \pmod{3},$$

and so

$$b^{560} = \left(b^2\right)^{280} \equiv 1 \pmod{3}.$$

Similarly,

$$b^{10} \equiv 1 \pmod{11},$$

and so

$$b^{560} = \left(b^{10}\right)^{56} \equiv 1 \pmod{11}.$$

Finally,

$$b^{16} \equiv 1 \pmod{17},$$

and so

$$b^{560} = \left(b^{16}\right)^{35} \equiv 1 \pmod{17}.$$

Since $b^{560} - 1$ is divisible by 3, 11, and 17, it is also divisible by their product, hence

$$b^{560} \equiv 1 \pmod{561}.$$

This proves that 561 is a pseudoprime to base b for every b such that $(b, n) = 1$.

A *Carmichael number* is a positive integer n such that n is composite but $b^{n-1} \equiv 1 \pmod{n}$ for every integer b relatively prime to n. Thus, 561 is a Carmichael number.

Exercises

1. Prove that 589 is composite by computing the least nonnegative residue of $2^{588} \pmod{589}$.
2. Let n be an odd integer, $n \geq 3$. Prove that there exists a nonnegative integer u such that $n + u^2 = (u+1)^2$. Prove that n is composite if and only if there exist nonnegative integers u and v such that $v > u + 1$ and $n + u^2 = v^2$. Use this method to factor 589.
3. Prove that 645 is a pseudoprime to base 2.
4. Prove that 1729 is a pseudoprime to bases 2, 3, and 5.
5. Prove that 1105 is a Carmichael number.
6. Let n be a product of distinct primes. Prove that if $p - 1$ divides $n - 1$ for every prime p that divides n, then n is a Carmichael number.
7. Prove that 6601 is a Carmichael number.

2.7 Public Key Cryptography

Cryptography is the art and science of sending secret messages. The message that we want to send is called the *plaintext*. The sender uses a *key* to encipher, or encrypt, it into *ciphertext*, and the ciphertext is transmitted

to the receiver, who uses another key to decipher, or decrypt, it back into plaintext. By writing letters and punctuation marks as numbers, we can assume that the plaintext is a positive integer P, and that it is encrypted as a different positive integer C. The problem is to invent keys that make it impossible or computationally infeasible for an enemy to decipher an intercepted message. *Cryptanalysis* is the art and science of deciphering an intercepted message without knowledge of the decrypting key.

Classically, cryptography uses secret keys that are known only to sender and receiver. If the enemy discovers the encrypting key and intercepts the ciphertext, then he might be able to compute the decrypting key and recover the plaintext.

Here is an example of a *secret key cryptosystem*. Let p be an odd prime, and let e be an integer such that $(e, p-1) = 1$. Suppose that the plaintext P is an integer such that $0 < P < p$. Let the ciphertext C be the least nonnegative residue of P^e modulo p, that is, we construct C by the rule

$$C \equiv P^e \pmod{p}$$

and

$$0 < C < p.$$

The encrypting key for this cipher consists of the prime number p and the integer e. To decrypt this cipher, we use elementary number theory. Since $(e, p-1) = 1$, there exists an integer d such that $ed \equiv 1 \pmod{p-1}$. It is easy to compute d. We can use the Euclidean algorithm, for example. The decrypting key consists of the prime p and the integer d. Since $ed = 1 + (p-1)k$ for some integer k, and since $P^{p-1} \equiv 1 \pmod{p}$ by Fermat's theorem, it follows that

$$C^d \equiv P^{ed} \equiv P^{1+(p-1)k} \equiv P(P^{p-1})^k \equiv P \pmod{p}.$$

Thus, we can decrypt the ciphertext C by computing the least nonnegative residue of C^d modulo p. An enemy who learns the encrypting key will break the cipher.

For example, if $p = 17$ and $e = 3$, then the plaintext $P = 10$ is encrypted as

$$P^3 = 10^3 \equiv 14 \pmod{17},$$

and so the ciphertext is $C = 14$. Since $3 \cdot 11 \equiv 1 \pmod{16}$, it follows that $d = 11$ is a decrypting key. We observe that

$$C^{11} = 14^{11} \equiv 10 = P \pmod{17}.$$

There is a more sophisticated idea in cryptography that produces secure ciphers even if the encrypting key is known. Indeed, the encrypting key can be made public, so that anyone can encrypt and send a message, but the decrypting key cannot be computed from knowledge of the encrypting key.

This is called a *public key cryptosystem*. Here is an example. We choose two different large primes p and q, and let

$$m = pq.$$

Since we know p and q, it is easy to calculate $\varphi(m) = (p-1)(q-1)$. Pick an integer e that is relatively prime to $\varphi(m)$. We publish the numbers m and e. The plaintext must be a positive integer P that is less than m and relatively prime to m If m is a large number, then almost all positive integers less than m are relatively prime to m (Exercise 4), so we can assume that $(P, m) = 1$. The ciphertext will be the unique integer C such that

$$C \equiv P^e \pmod{m}$$

and

$$0 < C < m.$$

It is important to note that we disclose neither $\varphi(m)$ nor the prime factors p and q of m. These are kept secret. However, since *we* know $\varphi(m)$, it is easy, by using the Euclidean algorithm, for example, to compute an integer d such that

$$ed \equiv 1 \pmod{\varphi(m)},$$

that is,

$$ed = 1 + \varphi(m)k$$

for some integer k. To decrypt the ciphertext C, we simply compute the least nonnegative residue of

$$C^d \pmod{m}.$$

Since $(P, m) = 1$, Euler's theorem tells us that

$$C^d \equiv P^{ed} \equiv P^{1+\varphi(m)k} \equiv P \pmod{m}.$$

The decryption key requires the integers d and m. It is not enough to know e and m. To compute d, one must know both e and $\varphi(m)$. Since $\varphi(m) = (p-1)(q-1)$, this requires a knowledge of the primes p and q such that $m = pq$, that is, we must be able to factor m. If the primes p and q are large (such as several thousand digits each), then it is impossible with state-of-the-art computer hardware and our current knowledge about factoring large numbers to find the prime factors of m in a reasonable time, for example, a million years. We know the prime factors p and q, and so we can compute $\varphi(m)$, but an opponent who wants to intercept and decrypt the message will fail, since he does not know the primes and cannot factor m. Indeed, the following result shows that knowing $\varphi(m)$ is equivalent to knowing the prime factors of m.

Theorem 2.19 *Let m be an integer that is the product of two prime numbers. The prime divisors of m are the roots of the quadratic equation*

$$x^2 - (m + 1 - \varphi(m))x + m = 0,$$

and so $\varphi(m)$ determines the prime factors of m.

Proof. If $m = pq$, then

$$\varphi(m) = (p-1)(q-1) = pq - p - q + 1 = m - p - \frac{m}{p} + 1,$$

and so

$$p - (m + 1 - \varphi(m)) + \frac{m}{p} = 0.$$

Equivalently, p and q are the solutions of the quadratic equation

$$x^2 - (m + 1 - \varphi(m))x + m = 0.$$

This completes the proof. □

For example, if $m = 221$ and $\varphi(m) = 192$, then the quadratic equation

$$x^2 - 30x + 221 = 0$$

has solutions $x = 13$ and $x = 17$, and $221 = 13 \cdot 17$.

This method, known as the *RSA cryptosystem*, is called a *public key cryptosystem*, since the encryption key is made available to everyone, and the encrypted message can be transmitted through public channels. Only the possessor of the prime factors of m can decrypt the message. RSA is simple, but useful, and is the basis of many commercially valuable cryptosystems.

Exercises

1. Consider the secret key cryptosystem constructed from the prime $p = 947$ and the encoding key $e = 167$. Encipher the plaintext $P = 2$. Find a decrypting key and decipher the ciphertext $C = 3$.

2. Consider the primes $p = 53$ and $q = 61$. Let $m = pq$. Prove that $e = 7$ is relatively prime to $\varphi(m)$. Find a positive integer d such that $ed \equiv 1 \pmod{\varphi(m)}$.

3. The integer 6059 is the product of two distinct primes, and $\varphi(6059) = 5904$. Use Theorem 2.19 to compute the prime divisors of 6059.

4. The probability that an integer chosen at random between 1 and n is relatively prime to n is $\varphi(n)/n$. Let $n = pq$, where p and q are distinct primes greater than x. Prove that the probability that a randomly chosen positive integer up to x is relatively prime to n is greater than $(1 - 1/x)^2$. If $x = 200$, this probability is greater than 0.99.

2.8 Notes

> Si numerus a numerorum b, c differentiam metitur, b et c *secundum a congrui* dicuntur, sin minus, *incongrui*: ipsum *a modulum* appellamus. Uterque numerorum b, c priori in casu alterius *residuum*, in posteriori vero *nonresiduum* vocatur.
>
> C. F. Gauss [37]

This is the first paragraph in the first section of Gauss's *Disquisitiones Arithmeticae*, a seminal book on number theory that was published in 1801. The translation, with slight changes in notation, is the first paragraph of this chapter. Gauss introduced the idea of congruence, and proved many of the results on congruences that we obtain in this book. This is classical mathematics that every student of mathematics should learn.

Carmichael conjectured in 1912 that the number of Carmichael numbers is infinite. Alford, Granville, and Pomerance [1] confirmed this in 1994. They proved that if $C(x)$ is the number of Carmichael numbers less than x, then $C(x) > x^{2/7}$ for all sufficiently large x. Erdős has made the stronger conjecture that for every $\varepsilon > 0$ there exists a number $x_0(\varepsilon)$ such that $C(x) > x^{1-\varepsilon}$ for all $x \geq x_0(\varepsilon)$. For an expository article on primality testing and Carmichael numbers, see Granville [40].

There is a vast literature on applications of number theory to cryptography, but it is hard to assign credit for discoveries in this field, because much of the research is carried out in secret at government agencies responsible for communications security, and not published in unclassified scientific journals. For example, the idea of public key cryptography first appeared in the public domain in work of Diffie, Hellman, and Merkle [26, 65] in 1976. The RSA cryptosystem was invented and published by Rivest, Shamir, and Adleman[123] in 1978. Singh [135] has reported, however, that both the concept of public key cryptography and the RSA cryptosystem were discovered earlier by three British government cryptographers, James Ellis, Clifford Cocks, and Malcolm Williamson, working at Government Communications Headquarters (GCHQ) in Cheltenham, England. It is possible that government cryptographers in other countries also independently discovered these methods.

Boneh [12] is a recent survey of the status of the RSA cryptosystem. In 1997, Shor [133] described an algorithm based on ideas from quantum mechanics that would factor large integers in "polynomial time," that is, much faster than is now possible with classical algorithms and computers. If it becomes possible to build quantum computers, then cryptography based on the difficulty of factoring large integers would become insecure and unreliable. For a review of classical computing, quantum computing, and Shor's factoring algorithm, see Manin [95]. Information on quantum

computing is available on the internet from the University of Oxford's Center for Quantum Computing (www.qubit.org).

A good text on number theoretic cryptography is Koblitz, *A Course in Number Theory and Cryptography* [83].

3

Primitive Roots and Quadratic Reciprocity

3.1 Polynomials and Primitive Roots

Let m be a positive integer greater than 1, and a an integer relatively prime to m. The *order of a modulo m*, denoted by $\text{ord}_m(a)$, is the smallest positive integer d such that $a^d \equiv 1 \pmod{m}$. By Theorem 2.14, $\text{ord}_m(a)$ is a divisor of the Euler phi function $\varphi(m)$. The order of a modulo m is also called the *exponent* of a modulo m.

We investigate the least nonnegative residues of the powers of a modulo m. For example, if $m = 7$ and $a = 2$, then

$$\begin{aligned} 2^0 &\equiv 1 \pmod{7}, \\ 2^1 &\equiv 2 \pmod{7}, \\ 2^2 &\equiv 4 \pmod{7}, \\ 2^3 &\equiv 1 \pmod{7}, \end{aligned}$$

and 2 has order 3 modulo 7. If $m = 7$ and $a = 3$, then

$$\begin{aligned} 3^0 &\equiv 1 \pmod{7}, \\ 3^1 &\equiv 3 \pmod{7}, \\ 3^2 &\equiv 2 \pmod{7}, \\ 3^3 &\equiv 6 \pmod{7}, \\ 3^4 &\equiv 4 \pmod{7}, \\ 3^5 &\equiv 5 \pmod{7}, \\ 3^6 &\equiv 1 \pmod{7}, \end{aligned}$$

and 3 has order 6 modulo 7. The powers of 3 form a reduced residue system modulo 7.

The integer a is called a *primitive root modulo* m if a has order $\varphi(m)$. In this case, the $\varphi(m)$ integers $1, a, a^2, \ldots, a^{\varphi(m)-1}$ are relatively prime to m and are pairwise incongruent modulo m. Thus, they form a reduced residue system modulo m. For example, 3 is a primitive root modulo 7. Similarly, 3 is a primitive root modulo 10, since $\varphi(10) = 4$ and

$$\begin{aligned}
3^0 &\equiv 1 \pmod{10}, \\
3^1 &\equiv 3 \pmod{10}, \\
3^2 &\equiv 9 \pmod{10}, \\
3^3 &\equiv 7 \pmod{10}, \\
3^4 &\equiv 1 \pmod{10}.
\end{aligned}$$

Some moduli do not have primitive roots. There is no primitive root modulo 8, for example, since $\varphi(8) = 4$, but

$$1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}, \tag{3.1}$$

and no integer has order 4 modulo 8.

In this section we prove that every prime p has a primitive root. In Section 3.2 we determine all composite moduli m for which there exist primitive roots.

We begin with some remarks about polynomials. Let R be a commutative ring with identity. A *polynomial with coefficients in* R is an expression of the form

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0,$$

where $a_0, a_1, \ldots, a_m \in R$. The element a_i is called the *coefficient* of the term x^i. The *degree* of the polynomial $f(x)$, denoted by $\deg(f)$, is the greatest integer n such that $a_n \neq 0$, and a_n is called the *leading coefficient.* If $\deg(f) = n$, we define $a_i = 0$ for $i > n$. Nonzero constant polynomials $f(x) = a_0 \neq 0$ have degree 0. The zero polynomial $f(x) = 0$ has no degree. A *monic polynomial* is a polynomial whose leading coefficient is 1.

We define addition and multiplication of polynomials in the usual way: If $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{j=0}^{m} b_j x^j$, then

$$(f+g)(x) = \sum_{k=0}^{\max(m,n)} (a_k + b_k) x^k$$

and

$$fg(x) = \sum_{k=0}^{mn} c_k x^k,$$

where

$$c_k = \sum_{\substack{i+j=k \\ 0 \leq i \leq n \\ 0 \leq j \leq m}} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i}.$$

With this addition and multiplication, the set $R[x]$ of all polynomials with coefficients in R is a commutative ring. Moreover,

$$\deg(f+g) \leq \max(\deg(f), \deg(g)).$$

If $f, g \in F[x]$ for some field F, then

$$\deg(fg) = \deg(f) + \deg(g),$$

and the leading coefficient of fg is $a_m b_n$.

For every $\alpha \in R$, the *evaluation map* $\Theta_\alpha : R[x] \to R$ defined by

$$\Theta_\alpha(f) = f(\alpha) = a_n \alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1 \alpha + a_0$$

is a ring homomorphism, that is, $(f+g)(\alpha) = f(\alpha) + g(\alpha)$ and $(fg)(\alpha) = f(\alpha)g(\alpha)$. The element α is called a *zero* or a *root* of the polynomial $f(x)$ if $\Theta_\alpha(f) = f(\alpha) = 0$.

We say that the polynomial $d(x)$ divides the polynomial $f(x)$ if there exists a polynomial $q(x)$ such that $f(x) = d(x)q(x)$.

Theorem 3.1 (Division algorithm for polynomials) *Let F be a field. If $f(x)$ and $d(x)$ are polynomials in $F[x]$ and if $d(x) \neq 0$, then there exist unique polynomials $q(x)$ and $r(x)$ such that $f(x) = d(x)q(x) + r(x)$ and either $r(x) = 0$ or the degree of $r(x)$ is strictly smaller than the degree of $d(x)$.*

Proof. Let $d(x) = b_m x^m + \cdots + b_1 x + b_0$, where $b_m \neq 0$ and $\deg(d) = m$. If $d(x)$ does not divide $f(x)$, then $f - dq \neq 0$ and $\deg(f - dq)$ is a nonnegative integer for every polynomial $q(x) \in F[x]$. Choose $q(x)$ such that $\ell = \deg(f - dq)$ is minimal, and let

$$r(x) = f(x) - d(x)q(x) = c_\ell x^\ell + \cdots + c_1 x + c_0 \in F[x],$$

where $c_\ell \neq 0$. We shall prove that $\ell < m$.

Since F is a field, $b_m^{-1} \in F$. If $\ell \geq m$, then

$$d(x) b_m^{-1} c_\ell x^{\ell - m}$$

is a polynomial of degree ℓ with leading coefficient c_ℓ. Then

$$Q(x) = q(x) + b_m^{-1} c_\ell x^{\ell - m} \in F[x],$$

and

$$\begin{aligned} R(x) &= f(x) - d(x)Q(x) \\ &= f(x) - d(x)\left(q(x) + b_m^{-1}c_\ell x^{\ell-m}\right) \\ &= r(x) - d(x)b_m^{-1}c_\ell x^{\ell-m} \end{aligned}$$

is a polynomial of degree at most $\ell - 1$. This contradicts the minimality of ℓ, and so $\ell < m$.

Next we prove that the polynomials $q(x)$ and $r(x)$ are unique. Suppose that

$$f(x) = d(x)q_1(x) + r_1(x) = d(x)q_2(x) + r_2(x),$$

where $q_1(x), q_2(x), r_1(x), r_2(x)$ are polynomials in $F[x]$ such that $r_i(x) = 0$ or $\deg(r_i) < \deg(d)$ for $i = 1, 2$. Then

$$d(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x).$$

If $q_1(x) \neq q_2(x)$, then

$$\deg(d) \leq \deg(d(q_1 - q_2)) = \deg(r_2 - r_1) < \deg(d),$$

which is absurd. Therefore, $q_1(x) = q_2(x)$, and so $r_1(x) = r_2(x)$. This completes the proof. □

Theorem 3.2 *Let $f(x) \in F[x]$, $f(x) \neq 0$, and let $N_0(f)$ denote the number of distinct zeros of $f(x)$ in F. Then $N_0(f)$ does not exceed the degree of $f(x)$, that is,*

$$N_0(f) \leq \deg(f).$$

Proof. We use the division algorithm for polynomials. Let $\alpha \in F$. Dividing $f(x)$ by $x - \alpha$, we obtain

$$f(x) = (x - \alpha)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < \deg(x-\alpha) = 1$, that is, $r(x) = r_0$ is a constant. Letting $x = \alpha$, we see that $r_0 = f(\alpha)$, and so

$$f(x) = (x - \alpha)q(x) + f(\alpha)$$

for every $\alpha \in F$. In particular, if α is a zero of $f(x)$, then $x - \alpha$ divides $f(x)$.

We prove the theorem by induction on $n = \deg(f)$. If $n = 0$, then $f(x)$ is a nonzero constant and $N_0(f) = 0$. If $n = 1$, then $f(x) = a_0 + a_1x$ with $a_1 \neq 0$, and $N_0(f) = 1$ since $f(x)$ has the unique zero $\alpha = -a_1^{-1}a_0$. Suppose that $n \geq 2$ and the theorem is true for all polynomials of degree

at most $n-1$. If $N_0(f) = 0$, we are done. If $N_0(f) \geq 1$, let $\alpha \in F$ be a zero of $f(x)$. Then

$$f(x) = (x - \alpha)q(x),$$

and

$$\deg(q) = n - 1.$$

If β is a zero of $f(x)$ and $\beta \neq \alpha$, then

$$0 = f(\beta) = (\beta - \alpha)q(\beta),$$

and so β is a zero of $q(x)$. Since $\deg(q) = n - 1$, the induction hypothesis implies that

$$N_0(f) \leq 1 + N_0(q) \leq 1 + \deg(q) = n.$$

This completes the proof. □

Theorem 3.3 *Let G be a finite subgroup of the multiplicative group of a field. Then G is cyclic.*

Proof. Let $|G| = m$. By Theorem 2.15, if $a \in G$, then the order of a is a divisor of m. For every divisor d of m, let $\psi(d)$ denote the number of elements of G of order d. If $\psi(d) \neq 0$, then there exists an element a of order d, and every element of the cyclic subgroup $\langle a \rangle$ generated by a satisfies $a^d = 1$. By Theorem 3.2, the polynomial $f(x) = x^d - 1 \in F[x]$ has at most d zeros, and so every zero of $f(x)$ belongs to the cyclic subgroup $\langle a \rangle$. In particular, every element of G of order d must belong to $\langle a \rangle$. By Theorem 2.18, a cyclic group of order d has exactly $\varphi(d)$ generators, where $\varphi(d)$ is the Euler phi function. Therefore, $\psi(d) = 0$ or $\psi(d) = \varphi(d)$ for every divisor d of m. Since every element of G has order d for some divisor d of m, it follows that

$$\sum_{d|m} \psi(d) = m.$$

By Theorem 2.8,

$$\sum_{d|m} \varphi(d) = m,$$

and so $\psi(d) = \varphi(d)$ for every divisor d of m. In particular, $\psi(m) = \varphi(m) \geq 1$, and so G is a cyclic group of order m. □

Theorem 3.4 *For every prime p, the multiplicative group of the finite field $\mathbf{Z}/p\mathbf{Z}$ is cyclic. This group has $\varphi(p-1)$ generators. Equivalently, for every prime p, there exist $\varphi(p-1)$ pairwise incongruent primitive roots modulo p.*

Proof. This follows immediately from Theorem 3.3, since $|(\mathbf{Z}/p\mathbf{Z})^\times| = p-1$. □

The following table lists the primitive roots for the first six primes.

p	$\varphi(p-1)$	primitive roots
2	1	1
3	1	2
5	2	$2,3$
7	2	$3,5$
11	4	$2,6,7,8$
13	4	$2,6,7,11$

Let p be a prime, and let g be a primitive root modulo p. If a is an integer not divisible by p, then there exists a unique integer k such that

$$a \equiv g^k \pmod{p}$$

and

$$k \in \{0, 1, \ldots, p-2\}.$$

This integer k is called the *index* of a with respect to the primitive root g, and is denoted by

$$k = \operatorname{ind}_g(a).$$

If k_1 and k_2 are any integers such that $k_1 \leq k_2$ and

$$a \equiv g^{k_1} \equiv g^{k_2} \pmod{p},$$

then

$$g^{k_2 - k_1} \equiv 1 \pmod{p},$$

and so

$$k_1 \equiv k_2 \pmod{p-1}.$$

If $a \equiv g^k \pmod{p}$ and $b \equiv g^\ell \pmod{p}$, then $ab \equiv g^k g^\ell = g^{k+\ell} \pmod{p}$, and so

$$\operatorname{ind}_g(ab) \equiv k + \ell \equiv \operatorname{ind}_g(a) + \operatorname{ind}_g(b) \pmod{p-1}.$$

The index map ind_g is also called the *discrete logarithm* to the base g modulo p.

For example, 2 is a primitive root modulo 13. Here is a table of $\operatorname{ind}_2(a)$ for $a = 1, \ldots, 12$:

a	$\operatorname{ind}_2(a)$	a	$\operatorname{ind}_2(a)$
1	0	7	11
2	1	8	3
3	4	9	8
4	2	10	10
5	9	11	7
6	5	12	6

By Theorem 2.18, if g is a primitive root modulo p, then g^k is a primitive root if and only if $(k, p-1) = 1$. For example, for $p = 13$ there are $\varphi(12) = 4$ integers k such that $0 \leq k \leq 11$ and $(k, 12) = 1$, namely, $k = 1, 5, 7, 11$, and so the four pairwise incongruent primitive roots modulo 13 are

$$\begin{aligned} 2^1 &\equiv 2 \pmod{13}, \\ 2^5 &\equiv 6 \pmod{13}, \\ 2^7 &\equiv 11 \pmod{13}, \\ 2^{11} &\equiv 7 \pmod{13}. \end{aligned}$$

Exercises

1. Find a primitive root modulo 23.
2. Find a primitive root modulo 41.
3. Prove that 2 is a primitive root modulo 101.
4. Compute $\text{ind}_2(27)$ modulo 101.
5. Compute $\text{ind}_2(19)$ modulo 101.
6. What is the order of 3 modulo 101? Is 3 a primitive root modulo 101?
7. Prove that 2 is a primitive root modulo 53.
8. Find all solutions of the congruence $2^x \equiv 22 \pmod{53}$.
9. Compute $\text{ind}_2(a)$ for all a not divisible by 53.
10. Let p be an odd prime, and let g be a primitive root modulo p. Prove that
$$(p-1)! \equiv g^{(p-2)(p-1)/2} \equiv -1 \pmod{p}.$$
Hint: Observe that
$$(p-1)! \equiv 1 \cdot g \cdot g^2 \cdots g^{p-2} \pmod{p}$$
and
$$\frac{(p-2)(p-1)}{2} = \frac{p(p-1)}{2} - (p-1).$$
This gives another proof of Wilson's theorem (Theorem 2.4).
11. Prove that if m has one primitive root, then there are exactly $\varphi(\varphi(m))$ pairwise incongruent primitive roots modulo m.
12. Let g and r be primitive roots modulo p. Prove that
$$\text{ind}_r(a) \equiv \text{ind}_g(a)\text{ind}_r(g) \pmod{p-1}$$
for every integer a relatively prime to p.

13. Let g be a primitive root modulo the odd prime p. Prove that $g^{(p-1)/2} \equiv -1 \pmod{p}$.

14. Let g be a primitive root modulo the odd prime p. Prove that $-g$ is a primitive root modulo p if and only if $p \equiv 1 \pmod{4}$.

15. Let $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{i=0}^{n} b_i x^i$ be polynomials with integer coefficients. Then $f(x)$ and $g(x)$ are called *congruent modulo* m, written $f(x) \equiv g(x) \pmod{m}$, if $a_i \equiv b_i \pmod{m}$ for $i = 0, 1, \ldots, n$. Let p be an odd prime, and let

$$f(x) = x^{p-1} - 1$$

and

$$g(x) = (x-1)(x-2)\cdots(x-(p-1)).$$

Prove the following statements:

(a) The polynomial $f(x) - g(x)$ has degree $p - 2$.

(b)

$$f(c) \equiv g(c) \equiv 0 \pmod{p} \quad \text{for } c = 1, 2, \ldots, p-1.$$

(c)

$$f(x) \equiv g(x) \pmod{p}.$$

Hint: Apply Theorem 3.2.

16. Prove that Exercise (15c) implies Wilson's theorem,

$$(p-1)! \equiv -1 \pmod{p}.$$

17. Prove that for every prime $p \geq 5$,

$$\sum_{1 \leq i < j \leq p-1} ij \equiv 0 \pmod{p}$$

and

$$\sum_{1 \leq i < j < k \leq p-1} ijk \equiv 0 \pmod{p}.$$

Hint: Exercise (15c).

18. Let R be a commutative ring with identity. An *ideal* of R is an additive subgroup $I \subseteq R$ such that, if $a \in I$ and $r \in R$, then $ar \in I$. Prove that if $I \neq \{0\}$ is an ideal of the polynomial ring $F[x]$, where F is a field, then there is a unique monic polynomial $d(x) \in I$ such that I consists of all multiples of $d(x)$, that is,

$$I = \{q(x)d(x) : q(x) \in F[x]\}.$$

Hint: If $I \neq \{0\}$, choose $d(x) \in I$ of minimal degree. The proof is similar to the proof of Theorem 1.3.

19. Prove that the intersection of a family of ideals is an ideal. This means that if $\{I_j\}_{j\in J}$ is a family of ideals in the ring R, then $I = \bigcap_{j\in J} I_j$ is an ideal in R.

20. Let $F[x]$ be the ring of polynomials with coefficients in the field F, and let $f(x), g(x) \in F[x]$. Prove that there exists a unique monic polynomial $d(x) \in F[x]$ such that $d(x)$ divides both $f(x)$ and $g(x)$, and every common divisor of $f(x)$ and $g(x)$ divides $d(x)$. The polynomial $d(x)$ is called the *greatest common divisor* of $f(x)$ and $g(x)$.

    Hint: Consider the ideal I generated by $f(x)$ and $g(x)$, that is, the set
    $$I = \{u(x)f(x) + v(x)g(x) : u(x), v(x) \in F[x]\},$$
    and apply Exercise 18.

21. Let $f : R \to S$ be a ring homomorphism. Prove that the kernel of f, that is, the set
    $$f^{-1}(0) = \{r \in R : f(r) = 0\}$$
    is an ideal of R.

22. Let $\alpha \in F$, and let $I(\alpha)$ be the set of all polynomials $f(x) \in F[x]$ such that $f(\alpha) = 0$. Prove that $I(\alpha)$ is the kernel of the evaluation map Θ_α and that $I(\alpha)$ is an ideal of $F[x]$.

23. Let A be a nonempty subset of F, and let $I(A)$ be the set of all polynomials $f(x) \in F[x]$ such that $f(\alpha) = 0$ for all $\alpha \in A$. Prove that $I(A)$ is an ideal of $F[x]$, and
    $$I(A) = \bigcap_{\alpha\in A} I(\alpha).$$

3.2 Primitive Roots to Composite Moduli

In the previous section we proved that primitive roots exist for every prime number. We also observed that primitive roots do not exist for every modulus. For example, congruence (3.1) shows that there is no primitive root modulo 8. The goal of this section is to prove that an integer $m \geq 2$ has a primitive root if and only if $m = 2, 4, p^k$, or $2p^k$, where p is an odd prime and k is a positive integer.

Theorem 3.5 *Let m be a positive integer that is not a power of 2. If m has a primitive root, then $m = p^k$ or $2p^k$, where p is an odd prime and k is a positive integer.*

Proof. Let a and m be integers such that $(a,m)=1$ and $m \geq 3$. Suppose that

$$m = m_1 m_2, \qquad \text{where } (m_1,m_2)=1 \text{ and } m_1 \geq 3,\ m_2 \geq 3. \tag{3.2}$$

Then $(a,m_1)=(a,m_2)=1$. The Euler phi function $\varphi(m)$ is even for $m \geq 3$ (Exercise 4 in Section 2.2). Let

$$n = \frac{\varphi(m)}{2} = \frac{\varphi(m_1)\varphi(m_2)}{2}.$$

By Euler's theorem,

$$a^{\varphi(m_1)} \equiv 1 \pmod{m_1},$$

and so

$$a^n = \left(a^{\varphi(m_1)}\right)^{\varphi(m_2)/2} \equiv 1 \pmod{m_1}.$$

Similarly,

$$a^n = \left(a^{\varphi(m_2)}\right)^{\varphi(m_1)/2} \equiv 1 \pmod{m_2}.$$

Since $(m_1,m_2)=1$ and $m=m_1m_2$, we have

$$a^n \equiv 1 \pmod{m},$$

and so the order of a modulo m is strictly smaller than $\varphi(m)$. Consequently, if we can factor m in the form (3.2), then there does not exist a primitive root modulo m. In particular, if m is divisible by two distinct odd primes, then m does not have a primitive root. Similarly, if $m = 2^\ell p^k$, where $\ell \geq 2$, then m does not have a primitive root. Therefore, the only moduli $m \neq 2^\ell$ for which primitive roots can exist are of the form $m = p^k$ or $m = 2p^k$ for some odd prime p. □

To prove the converse of Theorem 3.5, we use the following result about the exponential increase in the order of an integer modulo prime powers.

Theorem 3.6 *Let p be an odd prime, and let $a \neq \pm 1$ be an integer not divisible by p. Let d be the order of a modulo p. Let k_0 be the largest integer such that $a^d \equiv 1 \pmod{p^{k_0}}$. Then the order of a modulo p^k is d for $k = 1, \ldots, k_0$ and dp^{k-k_0} for $k \geq k_0$.*

Proof. There exists an integer u_0 such that

$$a^d = 1 + p^{k_0} u_0 \qquad \text{and} \qquad (u_0, p) = 1. \tag{3.3}$$

Let $1 \leq k \leq k_0$, and let e be the order of a modulo p^k. If $a^e \equiv 1 \pmod{p^k}$, then $a^e \equiv 1 \pmod{p}$, and so d divides e. By (3.3), we have $a^d \equiv 1 \pmod{p^k}$, and so e divides d. It follows that $e = d$.

Let $j \geq 0$. We shall show that there exists an integer u_j such that

$$a^{dp^j} = 1 + p^{j+k_0}u_j \qquad \text{and} \qquad (u_j, p) = 1. \tag{3.4}$$

The proof is by induction on j. The assertion is true for $j = 0$ by (3.3). Suppose we have (3.4) for some integer $j \geq 0$. By the binomial theorem, there exists an integer v_j such that

$$\begin{aligned} a^{dp^{j+1}} &= \left(1 + p^{j+k_0}u_j\right)^p \\ &= 1 + p^{j+1+k_0}u_j + \sum_{i=2}^{p}\binom{p}{i}p^{i(j+k_0)}u_j^i \\ &= 1 + p^{j+1+k_0}u_j + p^{j+2+k_0}v_j \\ &= 1 + p^{j+1+k_0}(u_j + pv_j) \\ &= 1 + p^{j+1+k_0}u_{j+1}, \end{aligned}$$

and the integer $u_{j+1} = u_j + pv_j$ is relatively prime to p. Thus, (3.4) holds for all $j \geq 0$.

Let $k \geq k_0 + 1$ and $j = k - k_0 \geq 1$. Suppose that the order of a modulo p^{k-1} is dp^{j-1}. Let e_k denote the order of a modulo p^k. The congruence

$$a^{e_k} \equiv 1 \pmod{p^k}$$

implies that

$$a^{e_k} \equiv 1 \pmod{p^{k-1}},$$

and so dp^{j-1} divides e_k. Since

$$a^{dp^{j-1}} = 1 + p^{k-1}u_{j-1} \not\equiv 1 \pmod{p^k},$$

it follows that dp^{j-1} is a proper divisor of e_k. On the other hand,

$$a^{dp^j} = 1 + p^k u_j \equiv 1 \pmod{p^k},$$

and so e_k divides dp^j. It follows that the order of a modulo p^k is exactly $e_k = dp^j = dp^{k-k_0}$. This completes the proof. □

Theorem 3.7 *Let p be an odd prime. If g is a primitive root modulo p, then either g or $g + p$ is a primitive root modulo p^k for all $k \geq 2$. If g is a primitive root modulo p^k and $g_1 \in \{g, g + p^k\}$ is odd, then g_1 is a primitive root modulo $2p^k$.*

Proof. Let g be a primitive root modulo p. The order of g modulo p is $p - 1$. Let k_0 be the largest integer such that p^{k_0} divides $g^{p-1} - 1$. By

Theorem 3.6, if $k_0 = 1$, then the order of g modulo p^k is $(p-1)p^{k-1} = \varphi(p^k)$, and g is a primitive root modulo p^k for all $k \geq 1$.

If $k_0 \geq 2$, then

$$g^{p-1} = 1 + p^2 v$$

for some integer v. By the binomial theorem,

$$\begin{aligned}(g+p)^{p-1} &= \sum_{i=0}^{p-1} \binom{p-1}{i} g^{p-1-i} p^i \\ &\equiv g^{p-1} + (p-1)g^{p-2}p \pmod{p^2} \\ &\equiv 1 + p^2 v + g^{p-2}p^2 - g^{p-2}p \pmod{p^2} \\ &\equiv 1 - g^{p-2}p \pmod{p^2} \\ &\not\equiv 1 \pmod{p^2}.\end{aligned}$$

Then $g + p$ is a primitive root modulo p such that

$$(g+p)^{p-1} = 1 + pu_0 \qquad \text{and} \qquad (u_0, p) = 1.$$

Therefore, $g + p$ is a primitive root modulo p^k for all $k \geq 1$.

Next we prove that primitive roots exist for all moduli of the form $2p^k$. If g is a primitive root modulo p^k, then $g + p^k$ is also a primitive root modulo p^k. Since p^k is odd, it follows that one of the two integers g and $g + p^k$ is odd, and the other is even. Let g_1 be the odd integer in the set $\{g, g + p^k\}$. Since $(g + p^k, p^k) = (g, p^k) = 1$, it follows that $(g_1, 2p^k) = 1$. The order of g_1 modulo $2p^k$ is not less than $\varphi(p^k)$, which is the order of g_1 modulo p^k, and not greater than $\varphi(2p^k)$. However, since p is an odd prime, we have

$$\varphi(2p^k) = \varphi(p^k),$$

and so g_1 has order $\varphi(2p^k)$ modulo $2p^k$, that is, g_1 is a primitive root modulo $2p^k$. This completes the proof. □

For example, 2 is a primitive root modulo 3. Since 3 is the greatest power of 3 that divides $2^2 - 1$, it follows that 2 is a primitive root modulo 3^k for all $k \geq 1$, and $2 + 3^k$ is a primitive root modulo $2 \cdot 3^k$ for all $k \geq 1$.

Finally, we consider primitive roots modulo powers of 2.

Theorem 3.8 *There exists a primitive root modulo $m = 2^k$ if and only if $m = 2$ or 4.*

Proof. We note that 1 is a primitive root modulo 2, and 3 is a primitive root modulo 4. We shall prove that if $k \geq 3$, then there is no primitive root modulo 2^k. Since $\varphi(2^k) = 2^{k-1}$, it suffices to show that

$$a^{2^{k-2}} \equiv 1 \pmod{2^k} \tag{3.5}$$

for a odd and $k \geq 3$. We do this by induction on k. The case $k = 3$ is congruence (3.1). Let $k \geq 3$, and suppose that (3.5) is true. Then

$$a^{2^{k-2}} - 1$$

is divisible by 2^k. Since a is odd, it follows that

$$a^{2^{k-2}} + 1$$

is even. Therefore,

$$a^{2^{k-1}} - 1 = \left(a^{2^{k-2}} - 1\right)\left(a^{2^{k-2}} + 1\right)$$

is divisible by 2^{k+1}, and so

$$a^{2^{k-1}} \equiv 1 \pmod{2^{k+1}}.$$

This completes the induction and the proof of theorem. □

Let $k \geq 3$. By Theorem 3.8, there is no primitive root modulo 2^k, that is, there does not exist an odd integer whose order modulo 2^k is 2^{k-1}. However, there do exist odd integers of order 2^{k-2} modulo 2^k.

Theorem 3.9 *For every positive integer k,*

$$5^{2^k} \equiv 1 + 3 \cdot 2^{k+2} \pmod{2^{k+4}}.$$

Proof. The proof is by induction on k. For $k = 1$ we have

$$5^{2^1} = 25 \equiv 1 + 3 \cdot 2^3 \pmod{2^5}.$$

Similarly, for $k = 2$ we have

$$5^{2^2} = 625 = 1 + 48 + 576 \equiv 1 + 3 \cdot 2^4 \pmod{2^6}.$$

If the theorem holds for $k \geq 1$, then there exists an integer u such that

$$5^{2^k} = 1 + 3 \cdot 2^{k+2} + 2^{k+4}u = 1 + 2^{k+2}(3 + 4u).$$

Since $2k + 4 \geq k + 5$, we have

$$\begin{aligned}
5^{2^{k+1}} &= \left(5^{2^k}\right)^2 \\
&= \left(1 + 2^{k+2}(3 + 4u)\right)^2 \\
&\equiv 1 + 2^{k+3}(3 + 4u) \pmod{2^{2k+4}} \\
&\equiv 1 + 3 \cdot 2^{k+3} \pmod{2^{k+5}}.
\end{aligned}$$

This completes the proof. □

Theorem 3.10 *If $k \geq 3$, then 5 has order 2^{k-2} modulo 2^k. If $a \equiv 1 \pmod 4$, then there exists a unique integer $i \in \{0, 1, \ldots, 2^{k-2} - 1\}$ such that*

$$a \equiv 5^i \pmod{2^k}.$$

If $a \equiv 3 \pmod 4$, then there exists a unique integer $i \in \{0, 1, \ldots, 2^{k-2}-1\}$ such that

$$a \equiv -5^i \pmod{2^k}.$$

Proof. In the case $k = 3$, we observe that 5 has order 2 modulo 8, and

$$\begin{aligned} 1 &\equiv 5^0 \pmod 8, \\ 3 &\equiv -5^1 \pmod 8, \\ 5 &\equiv 5^1 \pmod 8, \\ 7 &\equiv -5^0 \pmod 8. \end{aligned}$$

Let $k \geq 4$. By Theorem 3.9, we have

$$\begin{aligned} 5^{2^{k-2}} &\equiv 1 + 3 \cdot 2^k \pmod{2^{k+2}} \\ &\equiv 1 \pmod{2^k} \end{aligned}$$

and

$$\begin{aligned} 5^{2^{k-3}} &\equiv 1 + 3 \cdot 2^{k-1} \pmod{2^{k+1}} \\ &\equiv 1 + 3 \cdot 2^{k-1} \pmod{2^k} \\ &\not\equiv 1 \pmod{2^k}. \end{aligned}$$

Therefore, 5 has order exactly 2^{k-2} modulo 2^k, and so the integers 5^i are pairwise incongruent modulo 2^k for $i = 0, 1, \ldots, 2^{k-2} - 1$. Since $5^i \equiv 1 \pmod 4$ for all i, and since exactly half, that is, 2^{k-2}, of the 2^{k-1} odd numbers between 0 and 2^k are congruent to 1 modulo 4, it follows that the congruence

$$5^i \equiv a \pmod{2^k}$$

is solvable for every $a \equiv 1 \pmod 4$. If $a \equiv 3 \pmod 4$, then $-a \equiv 1 \pmod 4$ and so the congruence

$$-a \equiv 5^i \pmod{2^k},$$

or, equivalently,

$$a \equiv -5^i \pmod{2^k},$$

is solvable. This completes the proof. □

In algebraic language, Theorem 3.10 states that for all $k \geq 3$,

$$(\mathbf{Z}/2^k\mathbf{Z})^\times = \langle -1 \rangle \times \langle 5 \rangle \cong \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2^{k-2}\mathbf{Z},$$

where $\langle a \rangle$ denotes the cyclic subgroup of $(\mathbf{Z}/2^k\mathbf{Z})^\times$ generated by a for $a = -1$ and $a = 5$.

Exercises

1. Find an integer g that is a primitive root modulo 5^k for all $k \geq 1$. Find a primitive root modulo 10. Find a primitive root modulo 50.

2. For $k \geq 1$, let e_k be the order of 5 modulo 3^k. Prove that
$$e_k = 2 \cdot 3^{k-1}.$$

3. Prove that p divides the binomial coefficient $\binom{p}{i}$ for $i = 1, 2, \ldots, p-1$.

4. Prove that if g is a primitive root modulo p^2, then g is a primitive root modulo p^k for all $k \geq 2$.

5. Let p be an odd prime. Prove that
$$(1+px)^{p^k} \equiv 1 + p^{k+1}x \pmod{p^{k+2}}$$
for every integer x and every nonnegative integer k.

6. (Nathanson [100]; see also Wagstaff [151]) Let p be an odd prime, and let $a \neq \pm 1$ be an integer not divisible by p. Let d be the order of a modulo p, and let k_0 be the largest integer such that $a^d \equiv 1 \pmod{p^{k_0}}$. Prove that if $k \geq k_0$ is a solution of the exponential congruence
$$a^k \equiv 1 \pmod{p^k}, \tag{3.6}$$
then
$$\frac{p^k}{k} < \frac{a^d}{d},$$
and so congruence (3.6) has only finitely many solutions.

   Hint: Apply Theorem 3.6.

7. Use Exercise 6 to prove that the exponential congruence
$$9^k \equiv 1 \pmod{7^k}$$
has no solutions.

8. Find all solutions of the exponential congruence
$$17^k \equiv 1 \pmod{15^k}.$$

9. Find all solutions of the exponential congruence
$$3^k \equiv 1 \pmod{2^k}.$$

10. Let $\{x\}$ denote the fractional part of x. Compute

$$\left\{\left(\frac{3}{2}\right)^n\right\}$$

for $n = 1, \ldots, 10$. Let r_n be the least nonnegative residue of 3^n modulo 2^n. Show that

$$\left\{\left(\frac{3}{2}\right)^n\right\} = \frac{r_n}{3^n}.$$

Remark. It is an important unsolved problem in number theory to understand the distribution of the fractional parts of the powers of $3/2$ in the interval $[0, 1)$.

3.3 Power Residues

Let m, k, and a be integers such that $m \geq 2$, $k \geq 2$, and $(a, m) = 1$. We say that a is a *kth power residue modulo m* if there exists an integer x such that

$$x^k \equiv a \pmod{m}.$$

If this congruence has no solution, then a is called a *kth power nonresidue modulo m*.

Let $k = 2$ and $(a, m) = 1$. If the congruence $x^2 \equiv a \pmod{m}$ is solvable, then a is called a *quadratic residue modulo m*. Otherwise, a is called a *quadratic nonresidue modulo m*. For example, the quadratic residues modulo 7 are $1, 2$, and 4; the quadratic nonresidues are $3, 5$, and 6. The only quadratic residue modulo 8 is 1, and the quadratic nonresidues modulo 8 are $3, 5, 4$ and 7.

Let $k = 3$ and $(a, m) = 1$. If the congruence $x^3 \equiv a \pmod{m}$ is solvable, then a is called a *cubic residue modulo m*. Otherwise, a is called a *cubic nonresidue modulo m*. For example, the cubic residues modulo 7 are 1 and 6; the cubic nonresidues are 2, 3, 4, and 5. The cubic residues modulo 5 are 1, 2, 3, and 4; there are no cubic nonresidues modulo 5.

In this and the next two sections we investigate power residues modulo primes. In Section 3.6 we consider quadratic residues to composite moduli.

Theorem 3.11 *Let p be prime, $k \geq 2$, and $d = (k, p-1)$. Let a be an integer not divisible by p. Let g be a primitive root modulo p, Then a is a kth power residue modulo p if and only if*

$$ind_g(a) \equiv 0 \pmod{d}$$

if and only if

$$a^{(p-1)/d} \equiv 1 \pmod{p}.$$

If a is a kth power residue modulo p, then the congruence

$$x^k \equiv a \pmod{p} \tag{3.7}$$

has exactly d solutions that are pairwise incongruent modulo p. Moreover, there are exactly $(p-1)/d$ pairwise incongruent kth power residues modulo p.

Proof. Let $\ell = \text{ind}_g(a)$, where g is a primitive root modulo p. Congruence (3.7) is solvable if and only if there exists an integer y such that

$$g^y \equiv x \pmod{p}$$

and

$$g^{ky} \equiv x^k \equiv a \equiv g^\ell \pmod{p}.$$

This is equivalent to

$$ky \equiv \ell \pmod{p-1}. \tag{3.8}$$

This linear congruence in y has a solution if and only if

$$\text{ind}_g(a) = \ell \equiv 0 \pmod{d},$$

where $d = (k, p-1)$. Thus, the kth power residues modulo p are precisely the integers in the $(p-1)/d$ congruence classes $g^{id} + p\mathbf{Z}$ for $i = 0, 1, \ldots, (p-1)/d - 1$. Moreover,

$$a^{(p-1)/d} \equiv g^{(p-1)\ell/d} \equiv 1 \pmod{p}$$

if and only if

$$\frac{(p-1)\ell}{d} \equiv 0 \pmod{p-1}$$

if and only if

$$\text{ind}_g(a) = \ell \equiv 0 \pmod{d}.$$

Finally, if the linear congruence (3.8) is solvable, then by Theorem 2.2 it has exactly d solutions y that are pairwise incongruent modulo $p-1$, and so (3.7) has exactly d solutions $x = g^y$ that are pairwise incongruent modulo p. This completes the proof. □

For example, let $p = 19$ and $k = 3$. Then $d = (k, p-1) = (3, 18) = 3$. We can check that 2 is a primitive root modulo 19, and so a is a cubic residue modulo 19 if and only if 3 divides $\text{ind}_2(a)$. Since $-1 \equiv 2^9 \pmod{3}$ and $\text{ind}_2(-1) = 9$, it follows that -1 is a cubic residue modulo 19. The solutions of the congruence $x^3 \equiv -1 \pmod{19}$ are of the form $x \equiv 2^y \pmod{19}$, where $0 \leq y \leq 17$ and $3y \equiv 9 \pmod{18}$. Then $y \equiv 3 \pmod{6}$, and so

$y = 3, 9$, and 15. These give the following three cube roots of -1 modulo 19:

$$8 \equiv 2^3 \pmod{19},$$

$$18 \equiv 2^9 \pmod{19},$$

and

$$12 \equiv 2^{15} \pmod{19}.$$

Corollary 3.1 *Let p be an odd prime, and let $k \geq 2$ be an integer such that $(k, p-1) = 1$. If $(a, p) = 1$, then a is a kth power residue modulo p, and the congruence $x^k \equiv a \pmod{p}$ has a unique solution modulo p.*

Exercises

1. Find all cubic residues modulo 19.

2. Find all solutions of the congruence $x^3 \equiv 8 \pmod{19}$.

3. Define the map $f : (\mathbf{Z}/19\mathbf{Z})^\times \to (\mathbf{Z}/19\mathbf{Z})^\times$ by $f(x + 19\mathbf{Z}) = x^3 + 19\mathbf{Z}$. Prove that f is a homomorphism of the multiplicative group $(\mathbf{Z}/19\mathbf{Z})^\times$, and compute its kernel.

4. Find all fifth power residues modulo 11.

5. Find all sixth power residues modulo 11.

6. Define the map $f : (\mathbf{Z}/23\mathbf{Z})^\times \to (\mathbf{Z}/23\mathbf{Z})^\times$ by $f(x+23\mathbf{Z}) = x^3+23\mathbf{Z}$. Prove that f is a isomorphism of the multiplicative group $(\mathbf{Z}/23\mathbf{Z})^\times$, that is, prove that f is a homomorphism that is one-to-one and onto.

7. Let x_a be the least nonnegative integer such that $x_a^3 \equiv a \pmod{11}$. Compute x_a for $a = 1, 2, \ldots, 10$.

8. Prove that if $p \equiv 2 \pmod{3}$, then every integer not divisible by p is a cubic residue modulo p.

9. Prove that if $p \equiv 1 \pmod{6}$, then the product of the $(p-1)/3$ cubic residues modulo p is congruent to -1 modulo p.

3.4 Quadratic Residues

Let p be an odd prime and a an integer not divisible by p. Then a is called a *quadratic residue modulo p* if there exists an integer x such that

$$x^2 \equiv a \pmod{p}. \tag{3.9}$$

If this congruence has no solution, then a is called a *quadratic nonresidue modulo* p. Thus, an integer a is a quadratic residue modulo p if and only if $(a, p) = 1$ and a has a square root modulo p. By Theorem 3.11, exactly half the congruence classes relatively prime to p have square roots modulo p.

We define the *Legendre symbol* for the odd prime p as follows: For any integer a,

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } (a,p) = 1 \text{ and } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } (a,p) = 1 \text{ and } a \text{ is a quadratic nonresidue modulo } p, \\ 0 & \text{if } p \text{ divides } a. \end{cases}$$

The solvability of congruence (3.9) depends only on the congruence class of $a \pmod p$, that is,

$$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right) \quad \text{if } a \equiv b \pmod p,$$

and so the Legendre symbol is a well-defined function on the congruence classes $\mathbf{Z}/p\mathbf{Z}$.

We observe that if p is an odd prime, then, by Theorem 3.2, the only solutions of the congruence $x^2 \equiv 1 \pmod p$ are $x \equiv \pm 1 \pmod p$. Moreover, if $\varepsilon, \varepsilon' \in \{-1, 0, 1\}$ and $\varepsilon \equiv \varepsilon' \pmod p$, then p divides $\varepsilon - \varepsilon'$, and so $\varepsilon = \varepsilon'$. In particular, if $\left(\frac{a}{p}\right) \equiv \varepsilon \pmod p$, then $\left(\frac{a}{p}\right) = \varepsilon$.

Theorem 3.12 *Let p be an odd prime. For every integer a,*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod p.$$

Proof. If p divides a, then both sides of the congruence are 0. If p does not divide a, then, by Fermat's theorem,

$$\left(a^{(p-1)/2}\right)^2 \equiv a^{p-1} \equiv 1 \pmod p,$$

and so

$$a^{(p-1)/2} \equiv \pm 1 \pmod p.$$

Applying Theorem 3.11 with $k = 2$, we have

$$a^{(p-1)/2} \equiv 1 \pmod p \quad \text{if and only if} \quad \left(\frac{a}{p}\right) = 1,$$

and so

$$a^{(p-1)/2} \equiv -1 \pmod p \quad \text{if and only if} \quad \left(\frac{a}{p}\right) = -1.$$

This completes the proof. □

For example, 3 is a quadratic residue modulo the primes 11 and 13, and a quadratic nonresidue modulo the primes 17 and 19, because

$$\left(\frac{3}{11}\right) \equiv 3^5 \equiv 1 \pmod{11},$$

$$\left(\frac{3}{13}\right) \equiv 3^6 \equiv 1 \pmod{13},$$

$$\left(\frac{3}{17}\right) \equiv 3^8 \equiv -1 \pmod{17},$$

$$\left(\frac{3}{19}\right) \equiv 3^9 \equiv -1 \pmod{19}.$$

The next result states that the Legendre symbol is a completely multiplicative arithmetic function.

Theorem 3.13 *Let p be an odd prime, and let a and b be integers. Then*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Proof. If p divides a or b, then p divides ab, and

$$\left(\frac{ab}{p}\right) = 0 = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

If p does not divide ab, then, by Theorem 3.12,

$$\begin{aligned}\left(\frac{ab}{p}\right) &\equiv (ab)^{(p-1)/2} \pmod{p} \\ &\equiv a^{(p-1)/2}b^{(p-1)/2} \pmod{p} \\ &\equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod{p}.\end{aligned}$$

The result follows immediately from the observation that each side of this congruence is ± 1. □

Theorem 3.13 implies that the Legendre symbol $\left(\frac{\cdot}{p}\right)$ is completely determined by its values at -1, 2, and odd primes q. If a is an integer not divisible by p, then we can write

$$a = \pm 2^{r_0} q_1^{r_1} q_2^{r_2} \cdots q_k^{r_k},$$

where $q_1, \ldots, q_k$ are distinct odd primes not equal to p. Then

$$\left(\frac{a}{p}\right) = \left(\frac{\pm 1}{p}\right)\left(\frac{2}{p}\right)^{r_0}\left(\frac{q_1}{p}\right)^{r_1}\cdots\left(\frac{q_k}{p}\right)^{r_k}.$$

We shall first determine the set of primes p for which -1 is a quadratic residue. By the following result, this depends only on the congruence class of p modulo 4.

Theorem 3.14 *Let p be an odd prime number. Then*

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if} \quad p \equiv 1 \pmod 4, \\ -1 & \text{if} \quad p \equiv 3 \pmod 4). \end{cases}$$

Equivalently,

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}.$$

Proof. We observe that

$$(-1)^{(p-1)/2} = \begin{cases} 1 & \text{if} \quad p \equiv 1 \pmod 4, \\ -1 & \text{if} \quad p \equiv 3 \pmod 4. \end{cases}$$

Applying Theorem 3.12 with $a = -1$, we obtain

$$\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \pmod p.$$

Again, the theorem follows immediately from the observation that both sides of this congruence are ± 1. □

Let p be an odd prime, and let S be a set of $(p-1)/2$ integers. We call S a *Gaussian set* modulo p if $S \cup -S = S \cup \{-s : s \in S\}$ is a reduced system of residues modulo p. Equivalently, S is a Gaussian set if for every integer a not divisible by p, there exist $s \in S$ and $\varepsilon \in \{1, -1\}$ such that $a \equiv \varepsilon s \pmod p$. Moreover, s and ε are uniquely determined by a. For example, the sets $\{1, 2, \ldots, (p-1)/2\}$ and $\{2, 4, 6, \ldots, p-1\}$ are Gaussian sets modulo p for every odd prime p. If S is a Gaussian set, $s, s' \in S$, and $s \equiv \pm s' \pmod p$, then $s = s'$.

Theorem 3.15 (Gauss's lemma) *Let p be an odd prime, and a an integer not divisible by p. Let S be a Gaussian set modulo p. For every $s \in S$ there exist unique integers $u_a(s) \in S$ and $\varepsilon_a(s) \in \{1, -1\}$ such that*

$$as \equiv \varepsilon_a(s) u_a(s) \pmod p.$$

Moreover,

$$\left(\frac{a}{p}\right) = \prod_{s \in S} \varepsilon_a(s) = (-1)^m,$$

where m is the number of $s \in S$ such that $\varepsilon_a(s) = -1$.

Proof. Since S is a Gaussian set, for every $s \in S$ there exist unique integers $u_a(s) \in S$ and $\varepsilon_a(s) \in \{1, -1\}$ such that

$$as \equiv \varepsilon_a(s) u_a(s) \pmod{p}.$$

Let $s, s' \in S$. If $u_a(s) = u_a(s')$, then

$$\begin{aligned} as' &\equiv \varepsilon_a(s') u_a(s') \equiv \varepsilon_a(s') u_a(s) \pmod{p} \\ &\equiv \varepsilon_a(s') \varepsilon_a(s) \varepsilon_a(s) u_a(s) \pmod{p} \\ &\equiv \pm as \pmod{p}. \end{aligned}$$

Dividing by a, we obtain

$$s' \equiv \pm s \pmod{p},$$

and so $s' = s$. It follows that the map $u_a : S \to S$ is a permutation of S, and so

$$\prod_{s \in S} s = \prod_{s \in S} u_a(s).$$

Therefore,

$$\begin{aligned} a^{(p-1)/2} \prod_{s \in S} s &\equiv \prod_{s \in S} as \pmod{p} \\ &\equiv \prod_{s \in S} \varepsilon_a(s) u_a(s) \pmod{p} \\ &\equiv \prod_{s \in S} \varepsilon_a(s) \prod_{s \in S} u_a(s) \pmod{p} \\ &\equiv \prod_{s \in S} \varepsilon_a(s) \prod_{s \in S} s \pmod{p}. \end{aligned}$$

Dividing by $\prod_{s \in S} s$, we obtain

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \equiv \prod_{s \in S} \varepsilon_a(s) \pmod{p}.$$

The proof is completed by the observation that the right and left sides of this congruence are ± 1. □

We shall use Gauss's lemma to compute the Legendre symbol $\left(\frac{3}{11}\right)$. Let S be the Gaussian set $\{2, 4, 6, 8, 10\}$. We have

$$\begin{aligned} 3 \cdot 2 &\equiv 6 \pmod{11}, \\ 3 \cdot 4 &\equiv (-1)10 \pmod{11}, \\ 3 \cdot 6 &\equiv (-1)4 \pmod{11}, \\ 3 \cdot 8 &\equiv 2 \pmod{11}, \\ 3 \cdot 10 &\equiv 8 \pmod{11}. \end{aligned}$$

The number of $s \in S$ with $\varepsilon_3(s) = -1$ is $m = 2$, and so $\left(\frac{3}{11}\right) = (-1)^2 = 1$, that is, 3 is a quadratic residue modulo 11. Indeed,

$$5^2 \equiv 6^2 \equiv 3 \pmod{11},$$

and so 5 and 6 are the square roots of 3 modulo 11.

Theorem 3.16 *Let p be an odd prime. Then*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod{8}, \\ -1 & \text{if } p \equiv \pm 3 \pmod{8}. \end{cases}$$

Equivalently,

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

Proof. We apply Gauss's lemma (Theorem 3.15) to the Gaussian set $S = \{1, 2, 3, \ldots, (p-1)/2\}$. Then

$$\{2s : s \in S\} = \{2, 4, 6, \ldots, p-1\},$$

and

$$\left(\frac{2}{p}\right) = (-1)^m,$$

where m is the number of integers $s \in S$ such that $\varepsilon_2(s) = -1$. If $1 \leq 2s \leq (p-1)/2$, then $2s \in S$, and so $u_2(s) = 2s$ and $\varepsilon_2(s) = 1$. If $(p+1)/2 \leq 2s \leq p-1$, then $1 \leq p - 2s \leq (p-1)/2$, and so $p - 2s \in S$. Since

$$2s \equiv -(p-2s) \pmod{p},$$

it follows that $u_2(s) = p - 2s$ and $\varepsilon_2(s) = -1$. Therefore, m is the number of integers $s \in S$ such that $(p+1)/2 \leq 2s \leq p-1$, or, equivalently,

$$\frac{p+1}{4} \leq s \leq \frac{p-1}{2}. \tag{3.10}$$

Since every odd prime p is congruent to 1, 3, 5, or 7 modulo 8, there are four cases to consider.

(i) If $p \equiv 1 \pmod{8}$, then $p = 8k+1$, and $s \in S$ satisfies (3.10) if and only if

$$2k + \frac{1}{2} \leq s \leq 4k,$$

and so $m = 2k$ and $\left(\frac{2}{p}\right) = (-1)^{2k} = 1$.

(ii) If $p \equiv 3 \pmod{8}$, then $p = 8k+3$, and $s \in S$ satisfies (3.10) if and only if

$$2k + 1 \leq s \leq 4k + 1,$$

and so $m = 2k+1$ and $\left(\frac{2}{p}\right) = (-1)^{2k+1} = -1$.

(iii) If $p \equiv 5 \pmod 8$, then $p = 8k+5$, and $s \in S$ satisfies (3.10) if and only if
$$2k+1+\frac{1}{2} \leq s \leq 4k+2,$$
and so $m = 2k+1$ and $\left(\frac{2}{p}\right) = (-1)^{2k+1} = -1$.

(iv) If $p \equiv 7 \pmod 8$, then $p = 8k+7$, and $s \in S$ satisfies (3.10) if and only if
$$2k+2 \leq s \leq 4k+3,$$
and so $m = 2k+2$ and $\left(\frac{2}{p}\right) = (-1)^{2k+2} = 1$.

Finally, we observe that
$$\frac{p^2-1}{8} \equiv 0 \pmod 2 \quad \text{if } p \equiv 1 \text{ or } 7 \pmod 8$$
and
$$\frac{p^2-1}{8} \equiv 1 \pmod 2 \quad \text{if } p \equiv 3 \text{ or } 7 \pmod 8.$$
This completes the proof. □

Exercises

1. Find all solutions of the congruences $x^2 \equiv 2 \pmod{47}$ and $x^2 \equiv 2 \pmod{53}$.

2. Prove that $S = \{3, 4, 5, 9, 10\}$ is a Gaussian set modulo 11. Apply Gauss's lemma to this set to compute the Legendre symbols $\left(\frac{3}{11}\right)$ and $\left(\frac{7}{11}\right)$

3. Let p be an odd prime. Prove that $\{2, 4, 6, \ldots, p-1\}$ is a Gaussian set modulo p.

4. Use Theorem 3.14 and Theorem 3.16 to find all primes p for which -2 is a quadratic residue.

5. Use Gauss's lemma to find all primes p for which -2 is a quadratic residue.

6. Use Gauss's lemma to find all primes p for which 3 is a quadratic residue.

7. Find all primes p for which 4 is a quadratic residue.

8. Let p be an odd prime. Prove that the Legendre symbol is a homomorphism from the multiplicative group $(\mathbf{Z}/p\mathbf{Z})^\times$ into $\{\pm 1\}$. What is the kernel of this homomorphism?

9. For every odd prime p, define the *Mersenne number*

$$M_p = 2^p - 1.$$

A prime number of the form M_p is called a *Mersenne prime* (see Exercise 5 in Section 1.5).

Let q be a prime divisor of M_p.

(a) Prove that 2 has order p modulo q, and so p divides $q - 1$.

Hint: Fermat's theorem.

(b) Prove that p divides $(q-1)/2$, and so

$$q \equiv 1 \pmod{2p}$$

and

$$2^{(q-1)/2} \equiv 1 \pmod{q}.$$

Hint: Both p and q are odd.

(c) Prove that $\left(\frac{2}{q}\right) = 1$, and so $q \equiv \pm 1 \pmod{8}$.

10. For every positive integer n, define the *Fermat number*

$$F_n = 2^{2^n} + 1.$$

A prime number of the form F_n is called a *Fermat prime* (see Exercise 7 in Section 1.5).

Let $n \geq 2$, and let q be a prime divisor of F_n.

(a) Prove that 2 has order 2^{n+1} modulo q.

Hint: Exercise 8 in Section 2.5.

(b) Prove that

$$q \equiv 1 \pmod{2^{n+1}}.$$

(c) Prove that there exists an integer a such that

$$a^{2^{n+1}} \equiv -1 \pmod{q}.$$

Hint: Observe that $\left(\frac{2}{q}\right) = 1$, and so $2 \equiv a^2 \pmod{q}$.

(d) Prove that

$$q \equiv 1 \pmod{2^{n+2}}.$$

Remark. By Exercise 7 in Section 1.5, the Fermat number F_5 is divisible by the prime 641, and $641 \equiv 1 \pmod{2^7}$.

11. A *binary quadratic form* is a polynomial

$$f(x,y) = ax^2 + bxy + cy^2, \quad \text{where } a, b, c \text{ are integers.}$$

The *discriminant* of this form is the integer $d = b^2 - 4ac$. Show that

$$4af(x,y) = (2ax + by)^2 - dy^2.$$

12. Let p be an odd prime, and let $f(x,y) = ax^2 + bxy + cy^2$ be a binary quadratic form with $a \not\equiv 0 \pmod{p}$. We say that $f(x,y)$ has a nontrivial solution modulo p if there exist integers x and y not both divisible by p such that $f(x,y) \equiv 0 \pmod{p}$. Prove that $f(x,y)$ has a nontrivial solution modulo p if and only if either $d \equiv 0 \pmod{p}$ or d is a quadratic residue modulo p.

13. Prove that the binary quadratic form

$$f(x,y) = 2x^2 - 15xy + 27y^2$$

has a nontrivial solution modulo p for all primes p. Find a nontrivial solution of the congruence

$$f(x,y) \equiv 0 \pmod{11}.$$

14. Let p and q be distinct odd prime numbers. Prove that

$$\sum_{\substack{x_1+\cdots+x_q \equiv q \pmod{p} \\ 1 \leq x_i \leq p-1}} \left(\frac{x_1 \cdots x_q}{p}\right) \equiv 1 \pmod{q},$$

where the sum is over all ordered q-tuples of integers $(x_1, \ldots, x_q)$ such that $x_1 + \cdots + x_q \equiv q \pmod{p}$ and $1 \leq x_i \leq p-1$ for $i = 1, \ldots, q$.

Hint: If $qx \equiv q \pmod{p}$, then $x \equiv 1 \pmod{p}$. If the q-tuple $(x_1, \ldots, x_q)$ contains k distinct integers $y_1, \ldots, y_k$ such that integer y_j appears u_j times in the q-tuple, so that $\sum_{j=1}^k u_j y_j \equiv q \pmod{p}$ and $\sum_{j=1}^k u_j = q$, then the number of permutations of this q-tuple is the multinomial coefficient $\left(\frac{q}{u_1! \cdots u_k!}\right)$. Show that

$$\left(\frac{q}{u_1! \cdots u_k!}\right) \equiv 0 \pmod{q}.$$

3.5 Quadratic Reciprocity Law

Let p and q be distinct odd primes. If q is a quadratic residue modulo p, then the congruence

$$x^2 \equiv q \pmod{p}$$

is solvable. Similarly, if p is a quadratic residue modulo q, then the congruence

$$x^2 \equiv p \pmod{q}$$

is solvable. There is no obvious connection between these two congruences. One of the great discoveries of eighteenth-century mathematics is that there is, in fact, a subtle and powerful relation between them that depends only on the congruence classes of the primes p and q modulo 4. This is expressed in Gauss's celebrated *law of quadratic reciprocity.*

Theorem 3.17 (Quadratic reciprocity) *Let p and q be distinct odd primes. If $p \equiv 1 \pmod{4}$ or $q \equiv 1 \pmod{4}$, then p is a quadratic residue modulo q if and only if q is a quadratic residue modulo p. If $p \equiv q \equiv 3 \pmod{4}$, then p is a quadratic residue modulo q if and only if q is a quadratic nonresidue modulo p. Equivalently,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Proof. Let

$$S = \{1, 2, \ldots, (p-1)/2\}$$

and

$$T = \{1, 2, \ldots, (q-1)/2\}.$$

Then S is a Gaussian set for the prime p, and T is a Gaussian set for the prime q. Let

$$S \times T = \{(s,t) : s \in S, t \in T\}.$$

This is a rectangle of lattice points in $\mathbf{R}^2$ of cardinality

$$|S \times T| = \frac{p-1}{2}\frac{q-1}{2}.$$

We shall count the number m of lattice points (s,t) in this rectangle that lie in the strip defined by the inequality

$$1 \leq pt - qs \leq \frac{p-1}{2}. \tag{3.11}$$

(To understand this proof, it is helpful to choose small primes, for example, $p = 17$, $q = 13$, and draw pictures of the rectangle $S \times T$ and the regions defined by inequalities.)

If $s \in S$, $t_1, t_2 \in T$, and the lattice points (s, t_1) and (s, t_2) both satisfy (3.11), then

$$p|t_1 - t_2| = |(pt_1 - qs) - (pt_2 - qs)| < \frac{p-1}{2} < p,$$

and so $t_1 = t_2$. It follows that for every $s \in S$ there exists at most one $t \in T$ that satisfies (3.11). If this inequality holds for some $t \in T$, then $pt - qs = s' \in S$, and

$$qs \equiv -s' \pmod{p}.$$

Using the notation in Gauss's lemma (Theorem 3.15), we have $u_q(s) = s'$ and $\varepsilon_q(s) = -1$.

Conversely, if $s \in S$ and $\varepsilon_q(s) = -1$, then

$$qs \equiv -u_q(s) \pmod{p},$$

and there exists an integer t such that

$$qs = -u_q(s) + pt.$$

Since

$$0 < pt = qs + u_q(s) \leq \frac{q(p-1)}{2} + \frac{p-1}{2} = \frac{(q+1)(p-1)}{2},$$

it follows that

$$1 \leq t \leq \frac{(q+1)(p-1)}{2p} < \frac{q+1}{2}.$$

The prime q is odd, and so

$$1 \leq t \leq \frac{q-1}{2}.$$

Therefore, $t \in T$, and the lattice point $(s, t) \in S \times T$ satisfies inequality (3.11). Thus, the number m of lattice points $(s, t) \in S \times T$ that satisfy inequality (3.11) is equal to the number of $s \in S$ such that $\varepsilon_q(s) = -1$. By Gauss's lemma,

$$\left(\frac{q}{p}\right) = (-1)^m.$$

Similarly,

$$\left(\frac{p}{q}\right) = (-1)^n,$$

where n is the number of lattice points $(s, t) \in S \times T$ such that

$$1 \leq qs - pt \leq \frac{q-1}{2},$$

or, equivalently,

$$-\frac{q-1}{2} \leq pt - qs \leq -1. \tag{3.12}$$

Since $pt - qs \neq 0$ for all $s \in S$ and $t \in T$, it follows that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{m+n},$$

where $m + n$ is the number of lattice points $(s, t) \in S \times T$ such that

$$-\frac{q-1}{2} \leq pt - qs \leq \frac{p-1}{2}. \tag{3.13}$$

Let M denote the number of lattice points $(s, t) \in S \times T$ such that

$$pt - qs > \frac{p-1}{2}$$

and let N denote the number of lattice points $(s, t) \in S \times T$ such that

$$pt - qs < -\frac{q-1}{2}.$$

Then

$$m + n + M + N = |S \times T| = \frac{p-1}{2}\frac{q-1}{2}.$$

We define a map from the set $S \times T$ to itself by reflection:

$$(s, t) \mapsto (s', t'),$$

where

$$s' = \frac{p+1}{2} - s$$

and

$$t' = \frac{q+1}{2} - t.$$

This map is a bijection, since

$$\frac{p+1}{2} - s' = s$$

and

$$\frac{q+1}{2} - t' = t.$$

If $(s, t) \in S \times T$ and

$$pt - qs > \frac{p-1}{2},$$

then $(s',t') \in S \times T$ and

$$\begin{aligned} pt' - qs' &= p\left(\frac{q+1}{2} - t\right) - q\left(\frac{p+1}{2} - s\right) \\ &= \frac{p}{2} - pt - \frac{q}{2} + qs \\ &= -(pt - qs) + \frac{p-1}{2} - \frac{q-1}{2} \\ &< -\frac{q-1}{2}. \end{aligned}$$

Therefore, $M \leq N$. Similarly, if $(s,t) \in S \times T$ and

$$pt - qs < -\frac{q-1}{2},$$

then $(s',t') \in S \times T$ and

$$pt' - qs' > \frac{p-1}{2},$$

and so $M \geq N$. Therefore, $M = N$ and

$$\begin{aligned} \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) &= (-1)^{m+n} = (-1)^{m+n+2M} \\ &= (-1)^{m+n+M+N} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \end{aligned}$$

This completes the proof. □

The quadratic reciprocity law provides an effective method to calculate the value of the Legendre symbol. For example, since $7 \equiv 59 \equiv 3 \pmod 4$ and $59 \equiv 3 \pmod 7$, we have

$$\begin{aligned} \left(\frac{7}{59}\right) &= -\left(\frac{59}{7}\right) = -\left(\frac{3}{7}\right) \\ &= \left(\frac{7}{3}\right) = \left(\frac{1}{3}\right) \\ &= 1. \end{aligned}$$

Similarly, since $51 = 3 \cdot 17$ and $97 \equiv 17 \equiv 1 \pmod 4$, we have

$$\begin{aligned} \left(\frac{51}{97}\right) &= \left(\frac{3}{97}\right)\left(\frac{17}{97}\right) = \left(\frac{97}{3}\right)\left(\frac{97}{17}\right) \\ &= \left(\frac{1}{3}\right)\left(\frac{12}{17}\right) = \left(\frac{12}{17}\right) \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{4}{17}\right)\left(\frac{3}{17}\right) = \left(\frac{3}{17}\right) \\
&= \left(\frac{17}{3}\right) = \left(\frac{2}{3}\right) \\
&= -1.
\end{aligned}$$

Quadratic reciprocity also allows us to determine all primes p for which a given integer a is a quadratic residue. Here are some examples. If $a = 5$, then

$$\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 4 \pmod{5}, \\ -1 & \text{if } p \equiv 2, 3 \pmod{5}. \end{cases}$$

Let $a = 7$. If $p \equiv 1 \pmod{4}$, then

$$\left(\frac{7}{p}\right) = \left(\frac{p}{7}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 2, 4 \pmod{7}, \\ -1 & \text{if } p \equiv 3, 5, 6 \pmod{7}. \end{cases}$$

If $p \equiv 3 \pmod{4}$, then

$$\left(\frac{7}{p}\right) = -\left(\frac{p}{7}\right) = \begin{cases} 1 & \text{if } p \equiv 3, 5, 6 \pmod{7}, \\ -1 & \text{if } p \equiv 1, 2, 4 \pmod{7}. \end{cases}$$

Equivalently,

$$\left(\frac{7}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 3, 9, 19, 25, 27 \pmod{28}, \\ -1 & \text{if } p \equiv 5, 11, 13, 15, 17, 23 \pmod{28}. \end{cases}$$

Let $a = 35$. Then

$$\left(\frac{35}{p}\right) = 1$$

if and only if

$$p \equiv 1, 4 \pmod{5} \quad \text{and} \quad p \equiv 1, 3, 9, 19, 25, 27 \pmod{28}$$

or

$$p \equiv 2, 3 \pmod{5} \quad \text{and} \quad p \equiv 5, 11, 13, 15, 17, 23 \pmod{28}.$$

This is equivalent to a set of congruence classes modulo 140.

Exercises

1. Let $p = 11$ and $q = 7$. Using the notation in the proof of the law of quadratic reciprocity (Theorem 3.17), we have $m + n + M + N = |S \times T| = 15$. Compute the numbers m, n, M, and N. Check that $\left(\frac{7}{11}\right) = (-1)^m$ and $\left(\frac{11}{7}\right) = (-1)^n$.

2. Use quadratic reciprocity to compute $\left(\frac{7}{43}\right)$. Find an integer x such that $x^2 \equiv 7 \pmod{43}$.

3. Use quadratic reciprocity to compute $\left(\frac{19}{101}\right)$. Find an integer x such that $x^2 \equiv 19 \pmod{101}$.

4. Prove that the congruence

$$(x^2 - 2)(x^2 - 17)(x^2 - 34) \equiv 0 \pmod{p}$$

has a solution for every prime number p.

5. Use quaratic reciprocity to find all primes p for which -2 is a quadratic residue.

6. Use quaratic reciprocity to find all primes p for which 3 is a quadratic residue.

7. Find all primes for which -3 is a quadratic residue.

8. Find all primes for which 5 is a quadratic residue.

9. Find all primes for which -5 is a quadratic residue.

10. Find all primes p for which the binary quadratic form $f(x, y) = x^2 + xy + y^2$ has a nontrivial solution modulo p.

    Hint: Apply Exercise 11 in Section 3.4.

11. In Exercises 11–17 we derive properties of the *Jacobi symbol*, which is a generalization of the Legendre symbol to composite moduli. Let m be an odd positive integer, and let

$$m = \prod_{i=1}^{r} p_i^{k_i}$$

be the factorization of m into the product of powers of distinct prime numbers. For any nonzero integer a, we define the Jacobi symbol $\left(\frac{a}{m}\right)$ as follows:

$$\left(\frac{a}{m}\right) = \prod_{i=1}^{r} \left(\frac{a}{p_i}\right)^{k_i}.$$

(a) Prove that if $a \equiv b \pmod{m}$, then

$$\left(\frac{a}{m}\right) = \left(\frac{b}{m}\right).$$

(b) For any integers a and b, prove that

$$\left(\frac{ab}{m}\right) = \left(\frac{a}{m}\right)\left(\frac{b}{m}\right).$$

(c) Prove that $\left(\frac{a}{m}\right) = 0$ if and only if $(a, m) > 1$.

12. Compute the Jacobi symbol $\left(\frac{38}{165}\right)$.

13. Let m be an odd positive integer, and let $(a, m) = 1$. The integer a is called a quadratic residue modulo m if there exists an integer x such that
$$x^2 \equiv a \pmod{m}$$
and a quadratic nonresidue modulo m if this congruence has no solution. Prove that if $\left(\frac{a}{m}\right) = -1$, then a is a quadratic nonresidue modulo m. Prove that a is not necessarily a quadratic residue modulo m if $\left(\frac{a}{m}\right) = 1$.

    Hint: Consider $m = 21$ and $a = -1$.

14. Let $m = p^k$, where p is an odd prime and $k \geq 1$. Prove that
$$\frac{m-1}{2} \equiv \frac{k(p-1)}{2} \pmod{2}.$$

    Hint: Use the binomial theorem to expand $m = ((p-1)+1)^k$.

15. Let m be an odd positive integer with standard factorization $m = \prod_{i=1}^{r} p_i^{k_i}$. Prove that
$$\frac{m-1}{2} \equiv \sum_{i=1}^{r} \frac{k_i(p_i - 1)}{2} \pmod{2}.$$

    Hint: Use induction on r.

    Prove that
$$\left(\frac{-1}{m}\right) = (-1)^{(m-1)/2}.$$

16. Let m be an odd positive integer with standard factorization $m = \prod_{i=1}^{r} p_i^{k_i}$. Prove that
$$\frac{m^2-1}{8} \equiv \sum_{i=1}^{r} \frac{k_i(p_i^2 - 1)}{8} \pmod{8}$$
and
$$\left(\frac{2}{m}\right) = (-1)^{(m^2-1)/8}.$$

17. Let m and n be relatively prime odd positive integers with standard factorizations
$$m = \prod_{i=1}^{r} p_i^{k_i}$$

and

$$n = \prod_{j=1}^{s} q_j^{\ell_j}.$$

Prove that

$$\frac{m-1}{2}\frac{n-1}{2} \equiv \sum_{i=1}^{r}\sum_{j=1}^{s} k_i\ell_j \left(\frac{p_i-1}{2}\right)\left(\frac{q_j-1}{2}\right) \pmod 2$$

and

$$\left(\frac{n}{m}\right)\left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}.$$

3.6 Quadratic Residues to Composite Moduli

Let m be an odd positive integer and a an integer relatively prime to m. We shall prove that a is a quadratic residue modulo m if and only if a is a quadratic residue modulo p for every prime p that divides m. The Chinese remainder theorem (see Theorem 2.11) implies that it suffices to consider congruences modulo prime powers.

We begin with *Hensel's lemma*, an important result that gives a sufficient condition that a polynomial congruence solvable modulo a prime p will also be solvable modulo p^k for every positive integer k.

Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a polynomial with coefficients in a ring R. The *derivative* of $f(x)$ is the polynomial

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

If $f(x)$ is a polynomial of degree $n \geq 1$ with coefficients in the ring $\mathbf{Z}$, then the derivative $f'(x)$ has degree $n-1$ and leading coefficient na_n. For example, if $f(x) = x^3 - 5x + 1$, then $f'(x) = 3x^2 - 5$. Moreover,

$$\begin{aligned} f(x+h) &= (x+h)^3 - 5(x+h) + 1 \\ &= (x^3 + 3x^2h + 3xh^2 + h^3) - (5x + 5h) + 1 \\ &= (x^3 - 5x + 1) + (3x^2 - 5)h + (3x + h)h^2 \\ &= f(x) + f'(x)h + r(x,h)h^2, \end{aligned}$$

where $r(x,h) = 3x + h$.

Theorem 3.18 *Let R be a ring and $f(x) = \sum_{i=0}^{n} a_i x^i$ a polynomial with coefficients in R. Then*

$$f(x+h) = f(x) + f'(x)h + r(x,h)h^2.$$

where $r(x,h)$ is a polynomial in the two variables x and h with coefficients in R.

Proof. This is a standard calculation. Expanding $f(x+h)$ by the binomial theorem, we obtain

$$\begin{aligned} f(x+h) &= \sum_{i=0}^{n} a_i(x+h)^i \\ &= \sum_{i=0}^{n} a_i \sum_{j=0}^{i} \binom{i}{j} x^{i-j}h^j \\ &= \sum_{j=0}^{n}\sum_{i=j}^{n} \binom{i}{j} a_i x^{i-j}h^j \\ &= \sum_{i=0}^{n} a_i x^i + \sum_{i=1}^{n} i a_i x^{i-1}h + \sum_{j=2}^{n}\sum_{i=j}^{n} \binom{i}{j} a_i x^{i-j}h^j \\ &= f(x) + f'(x)h + r(x,h)h^2, \end{aligned}$$

where

$$r(x,h) = \sum_{j=2}^{n}\sum_{i=j}^{n} \binom{i}{j} a_i x^{i-j}h^{j-2}$$

is a polynomial in x and h with coefficients in R. □

Theorem 3.19 (Hensel's lemma) *Let p be prime, and let $f(x)$ be a polynomial of degree n with integer coefficients and leading coefficient not divisible by p. If there exists an integer x_1 such that*

$$f(x_1) \equiv 0 \pmod{p}$$

and

$$f'(x_1) \not\equiv 0 \pmod{p},$$

then for every $k \geq 2$ there exists an integer x_k such that

$$f(x_k) \equiv 0 \pmod{p^k} \tag{3.14}$$

and

$$x_k \equiv x_{k-1} \pmod{p^{k-1}}. \tag{3.15}$$

Proof. The proof is by induction on k. We begin by constructing x_2. There exist integers u_1 and v_1 such that $f(x_1) = u_1 p$ and $f'(x_1) = v_1 \not\equiv 0 \pmod{p}$. We shall prove that there exists an integer y_1 such that $f(x_1 + y_1 p) \equiv 0 \pmod{p^2}$.

By Theorem 3.18, there exists a polynomial $r(x,h)$ with integer coefficients such that

$$\begin{aligned} f(x_1 + y_1 p) &= f(x_1) + f'(x_1)y_1 p + r(x_1, y_1 p)p^2 \\ &= u_1 p + v_1 y_1 p + r(x_1, y_1 p)p^2 \\ &\equiv u_1 p + v_1 y_1 p \pmod{p^2}. \end{aligned}$$

Therefore, there exists an integer y_1 such that

$$f(x_1 + y_1 p) \equiv 0 \pmod{p^2}$$

if and only if the linear congruence

$$v_1 y \equiv -u_1 \pmod{p}$$

is solvable. We see that this congruence does have a solution y_1 because $(v_1, p) = 1$. Let

$$x_2 = x_1 + y_1 p.$$

Then

$$f(x_2) \equiv 0 \pmod{p} \quad \text{and} \quad x_2 \equiv x_1 \pmod{p}.$$

Let $k \geq 3$, and assume that we have constructed integers $x_2, \ldots, x_{k-1}$ such that

$$f(x_i) \equiv 0 \pmod{p^i} \quad \text{and} \quad x_i \equiv x_{i-1} \pmod{p^{i-1}}$$

for $i = 2, \ldots, k-1$. There exists an integer u_{k-1} such that

$$f(x_{k-1}) = u_{k-1} p^{k-1}.$$

Let $f'(x_{k-1}) = v_{k-1}$. Since $x_{k-1} \equiv x_1 \pmod{p}$, it follows that

$$v_{k-1} = f'(x_{k-1}) \equiv f'(x_1) \not\equiv 0 \pmod{p}.$$

Applying Theorem 3.18 with $t = x_{k-1}$ and $h = y_{k-1}p^{k-1}$, we obtain

$$\begin{aligned} & f\left(x_{k-1} + y_{k-1}p^{k-1}\right) \\ &= f\left(x_{k-1}\right) + f'(x_{k-1})y_{k-1}p^{k-1} + r(x_{k-1}, y_{k-1}p^{k-1})y_{k-1}^2 p^{2k-2} \\ &\equiv u_{k-1}p^{k-1} + v_{k-1}y_{k-1}p^{k-1} \pmod{p^k}. \end{aligned}$$

It follows that

$$f\left(x_{k-1} + y_{k-1}p^{k-1}\right) \equiv 0 \pmod{p^k}$$

if and only if there exists an integer y_{k-1} such that

$$v_{k-1} y_{k-1} \equiv -u_{k-1} \pmod{p}.$$

This last congruence is solvable, since $(v_{k-1}, p) = 1$, and the integer $x_k = x_{k-1} + y_{k-1}p^{k-1}$ satisfies conditions (3.14) and (3.15). □

Theorem 3.20 *Let p be an odd prime, and let a be an integer not divisible by p. If a is a quadratic residue modulo p, then a is a quadratic residue modulo p^k for every $k \geq 1$.*

Proof. Consider the polynomial $f(x) = x^2 - a$ and its derivative $f'(x) = 2x$. If a is a quadratic residue modulo p, then there exists an integer x_1 such that $x_1 \not\equiv 0 \pmod{p}$ and $x_1^2 \equiv a \pmod{p}$. Then $f(x_1) \equiv 0 \pmod{p}$ and $f'(x_1) \not\equiv 0 \pmod{p}$. By Hensel's lemma, the polynomial congruence $f(x) \equiv 0 \pmod{p^k}$ is solvable for every $k \geq 1$, and so a is a quadratic residue modulo p^k for every $k \geq 1$. □

Exercises

1. Let $x_1 = 3$. Costruct integers x_k such that $x_k^2 \equiv 2 \pmod{7^k}$ and $x_k \equiv x_{k-1} \pmod{7^{k-1}}$ for $k = 2, 3, 4$.

2. Let p be a prime, $p \neq 3$, and let a be an integer not divisible by p. Prove that if a is a cubic residue modulo p, then a is a cubic residue modulo p^k for every $k \geq 1$.

3. Denote the derivative of the polynomial $f(x)$ by $D(f)(x) = f'(x)$. We define

$$\begin{aligned} D^{(0)}(f)(x) &= f(x), \\ D^{(k)}(f)(x) &= D\left(D^{(k-1)}(f)\right)(x) \quad \text{for } k \geq 1. \end{aligned}$$

   The polynomial $D^{(k)}(f)$ is called the *kth derivative of f*. Prove that if $f(x)$ is a polynomial with integer coefficients, then $D^{(k)}(f)(x) = 0$ if and only if the degree of $f(x)$ is at most $k - 1$.

4. Let $f(x)$ and $g(x)$ be polynomials. Prove the *Leibniz formula*

$$D(f \cdot g)(x) = f(x) \cdot D(g)(x) + D(f)(x) \cdot g(x).$$

5. Let $f(x)$ be a polynomial of degree n. Prove *Taylor's formula*

$$f(x+h) = \sum_{k=0}^{n} \frac{D^{(k)}(x)}{k!} h^k.$$

6. This exercise generalizes Hensel's lemma (Theorem 3.19). Let p be a prime, and $f(x)$ a polynomial of degree n with integer coefficients and leading coefficient not divisible by p. Let ℓ be a nonnegative integer. If there exists an integer x_1 such that

$$\begin{aligned} f(x_1) &\equiv 0 \pmod{p^{2\ell+1}}, \\ f'(x_1) &\equiv 0 \pmod{p^{\ell}}, \end{aligned}$$

and

$$f'(x_1) \not\equiv 0 \pmod{p^{\ell+1}},$$

then for every $k \geq 2$ there exists an integer x_k such that

$$f(x_k) \equiv 0 \pmod{p^{2\ell+k}}$$

and

$$x_k \equiv x_{k-1} \pmod{p^{\ell+k-1}}.$$

Hint: Prove by induction on k. To begin the induction, find an integer y_1 such that $f(x_1 + y_1 p^{\ell+1}) \equiv 0 \pmod{p^{2\ell+2}}$ and let $x_2 = x_1 + y_1 p^{\ell+1}$.

3.7 Notes

Primitive roots and quadratic reciprocity are classical topics in number theory and a standard part of an introductory course in the subject.

There are still many simple questions about primitive roots that we cannot answer. For example, we cannot determine the prime numbers for which 2 is a primitive root. We do not even know if the number of such primes is finite or infinite. Gauss conjectured that 10 is a primitive root for infinitely many primes. This would imply, by Exercise 9 in Section 2.5, there are infinitely many primes p such that the decimal expansion of the fraction $1/p$ has period $p-1$. We do not, in fact, know even one integer that is a primitive root for infinitely many primes. There is an amazing result due to Gupta and Murty [44] and Heath-Brown [62] that states that every prime number, with at most two exceptions, is a primitive root for infinitely many primes. It follows that at least one of the numbers $2, 3$, and 5 is a primitive root for infinitely many primes, but we do not know which one.

Let a be an integer that is not a square and $a \neq -1$. A conjecture of Artin [5, page viii] states that there exist infinitely many primes for which a is a primitive root. Moreover, Artin has a conjectured density for the set of primes for which a is a primitive root. Murty [98] is a nice survey paper of Artin's conjecture and its generalizations. Erdős asked the following: For every sufficiently large prime p, does there exist a prime $q < p$ such that q is a primitive root modulo p?

4

Fourier Analysis on Finite Abelian Groups

4.1 The Structure of Finite Abelian Groups

This chapter introduces analysis on finite abelian groups and their characters. We begin by using elementary number theory to determine the structure of finite abelian groups.

Let G be an abelian group, written additively, and let $A_1, \ldots, A_k$ be subsets of G. The *sum* of these sets is the set

$$A_1 + \cdots + A_k = \{a_1 + \cdots + a_k : a_i \in A_i \text{ for } i = 1, \ldots, k\}.$$

If $G_1, \ldots, G_k$ are subgroups of G, then the sumset $G_1 + \cdots + G_k$ is a subgroup of G (Exercise 2). We say that G is the *direct sum* of the subgroups $G_1, \ldots, G_k$, written $G = G_1 \oplus \cdots \oplus G_k$, if every element $g \in G$ can be written uniquely in the form $g = g_1 + \cdots + g_k$, where $g_i \in G_i$ for $i = 1, \ldots, k$. If $G = G_1 \oplus \cdots \oplus G_k$, then $|G| = |G_1| \cdots |G_k|$ (Exercise 3).

The order of an element g in an additive group is the smallest positive integer d such that $dg = 0$. By Theorem 2.16, the order of an element of a finite group divides the order of the group.

Let p be a prime number. A *p-group* is a group each of whose elements has an order that is a power of p. For every prime number p, let $G(p)$ denote the set of all elements of G whose order is a power of p. Then $G(p)$ is a subgroup of the abelian group G (Exercise 6).

Theorem 4.1 *Let G be a finite abelian group, written additively, and let $|G| = m$. For every prime number p, let $G(p)$ be the set of all elements*

$g \in G$ whose order is a power of p. Then

$$G = \bigoplus_{p|m} G(p).$$

Proof. Let $m = \prod_{i=1}^{k} p_i^{r_i}$ be the standard factorization of m, and let $m_i = mp_i^{-r_i}$ for $i = 1, \ldots, k$. Then $(m_1, \ldots, m_k) = 1$ by Exercise 15 in Section 1.4, and so there exist integers $u_1, \ldots, u_k$ such that

$$m_1 u_1 + \cdots + m_k u_k = 1.$$

Let $g \in G$, and define $g_i = m_i u_i g \in G$ for $i = 1, \ldots, k$. Since $p_i^{r_i} g_i = m u_i g = 0$, it follows that $g_i \in G(p)$. Moreover,

$$\begin{aligned} g &= (m_1 u_1 + \cdots + m_k u_k) g = m_1 u_1 g + \cdots + m_k u_k g \\ &= g_1 + \cdots + g_k \in G(p_1) + \cdots + G(p_k), \end{aligned}$$

and so

$$G = G(p_1) + \cdots + G(p_k).$$

Suppose that

$$g_1 + \cdots + g_k = 0,$$

where $g_i \in G(p_i)$ for $i = 1, \ldots, k$. There exist nonnegative integers $r_1, \ldots, r_k$ such that g_i has order $p_i^{r_i}$ for $i = 1, \ldots, k$. Let

$$d_j = \prod_{\substack{i=1 \\ i \neq j}}^{k} p_i^{r_i}.$$

If $g_j \neq 0$, then $d_j g_j \neq 0$. Since $d_j g_i = 0$ for $i = 1, \ldots, k$, $i \neq j$, it follows that

$$0 = d_j (g_1 + \cdots + g_k) = d_j g_j,$$

and so $g_j = 0$ for all $j = 1, \ldots, k$. Thus, 0 has no nontrivial representation in $G = G(p_1) + \cdots + G(p_k)$. By Exercise 4, we conclude that G is the direct sum of the subgroups $G(p_i)$. □

Lemma 4.1 *Let G be a finite abelian p-group. Let $g_1 \in G$ be an element of maximum order p^{r_1}, and let $G_1 = \langle g_1 \rangle$ be the cyclic subgroup generated by g_1. Consider the quotient group G/G_1. Let $h \in G$. If $h + G_1 \in G/G_1$ has order p^r, then there exists an element $g \in G$ such that $g + G_1 = h + G_1$ and g has order p^r in G.*

Proof. If $h + G_1$ has order p^r in G/G_1, then the order of h in G is at most p^{r_1} (since p^{r_1} is the maximum order in G) and at least p^r (by

Exercise 7). Since $G_1 = p^r(h + G_1) = p^r h + G_1$, it follows that $p^r h \in G_1$, and so $p^r h = ug_1$ for some positive integer $u \leq p^{r_1}$ (since g_1 has order p^{r_1}). Write $u = p^s v$, where $(p, v) = 1$ and $0 \leq s \leq r_1$. Then vg_1 also has order p^{r_1}, and so $p^s vg_1$ has order $p^{r_1 - s}$ in G. Then $p^r h = p^s vg_1$ has order p^{r_1-s} in G, and so h has order $p^{r_1+r-s} \leq p^{r_1}$. It follows that $r \leq s$, and

$$p^r h = p^s vg_1 = p^r(p^{s-r}vg_1) = p^r g_1',$$

where

$$g_1' = p^{s-r}vg_1 \in G_1.$$

Let

$$g = h - g_1'.$$

Then

$$g + G_1 = h + G_1.$$

Moreover, $p^r g = p^r h - p^r g_1' = 0$, and so the order of g is at most p^r. On the other hand, $g + G_1$ has order p^r in the quotient group G/G_1, and so the order of g is at least p^r. Therefore, g has order p^r. □

Theorem 4.2 *Every finite abelian p-group is a direct sum of cyclic groups.*

Proof. The proof is by induction on the cardinality of G. Let G be a finite abelian p-group. If G is cyclic, we are done. If G is not cyclic, let $g_1 \in G$ be an element of maximum order p^{r_1}, and let G_1 be the cyclic subgroup generated by g_1. The quotient group G/G_1 is a finite abelian p-group, and

$$1 < |G/G_1| = \frac{|G|}{p^{r_1}} < |G|.$$

Therefore, the induction hypothesis holds for G/G_1, and so

$$G/G_1 = H_2 \oplus \cdots \oplus H_k,$$

where H_i is a cyclic subgroup of G/G_1 of order p^{r_i} for $i = 2, \ldots, k$. Moreover,

$$\frac{|G|}{p^{r_1}} = |G/G_1| = \prod_{i=2}^{k} p^{r_i}.$$

By Lemma 4.1, for each $i = 2, \ldots, k$ there exists an element $g_i \in G$ such that $g_i + G_1$ generates H_i and g_i has order p^{r_i} in G. Let G_i be the cyclic subgroup of G generated by g_i. Then $|G_i| = p^{r_i}$ for $i = 1, \ldots, k$. We shall prove that $G = G_1 \oplus \cdots \oplus G_k$.

We begin by showing that $G = G_1 + \cdots + G_k$. If $g \in G$, then $g + G_1 \in G/G_1$, and there exist integers $u_2, \ldots, u_k$ such that

$$0 \leq u_i \leq p^{r_i} - 1 \quad \text{for } i = 2, \ldots, k$$

and

$$g + G_1 = u_2(g_2 + G_1) \oplus \cdots \oplus u_k(g_k + G_1) = (u_2 g_2 + \cdots + u_k g_k) + G_1.$$

It follows that

$$g - (u_2 g_2 + \cdots + u_k g_k) = u_1 g_1 \in G_1$$

for some integer u_1 such that

$$0 \leq u_1 \leq p^{r_1} - 1,$$

and so

$$g = u_1 g_1 + u_2 g_2 + \cdots + u_k g_k \in G_1 + \cdots + G_k.$$

Therefore, $G = G_1 + \cdots + G_k$. Since

$$|G| = |G_1 + \cdots + G_k| \leq |G_1| \cdots |G_k| = \prod_{i=1}^{k} p^{r_i} = |G|,$$

it follows that every element of G has a unique representation as an element in the sumset $G_1 + \cdots + G_k$, and so $G = G_1 \oplus + \cdots + \oplus G_k$. This completes the proof. □

Theorem 4.3 *Every finite abelian group is a direct sum of cyclic groups.*

Proof. This follows immediately from Theorem 4.1 and Theorem 4.2. □

Let $G_1, \ldots, G_k$ be abelian groups, written additively. Their *direct product* is the group

$$G_1 \times \cdots \times G_k = \{(g_1, \ldots, g_k) : g_i \in G_i \text{ for } i = 1, \ldots, k\},$$

with addition defined by

$$(g_1, \ldots, g_k) + (g_1', \ldots, g_k') = (g_1 + g_1', \ldots, g_k + g_k').$$

If $G_1, \ldots, G_k$ are subgroups of an abelian group G and if $G = G_1 \oplus \cdots \oplus G_k$, then $G \cong G_1 \times \cdots \times G_k$ (Exercise 5).

Let $G_1, \ldots, G_k$ be abelian groups, written multiplicatively. Their *direct product* is the group $G_1 \times \cdots \times G_k$ consisting of all k-tuples $(g_1, \ldots, g_k)$ with $g_i \in G_i$ for $i = 1, \ldots, k$ and multiplication defined coordinate-wise by $(g_1, \ldots, g_k)(g_1', \ldots, g_k') = (g_1 g_1', \ldots, g_k g_k')$.

Exercises

1. Let $G = \mathbf{Z}/12\mathbf{Z}$ be the additive group of congruence classes modulo 12. Compute $G(2)$ and $G(3)$ and show explicitly that $G(2) \cong \mathbf{Z}/4\mathbf{Z}$, $G(3) \cong \mathbf{Z}/3\mathbf{Z}$, and
$$\mathbf{Z}/12\mathbf{Z} \cong \mathbf{Z}/4\mathbf{Z} \oplus \mathbf{Z}/3\mathbf{Z}.$$

2. Let G be an abelian group, written additively, and let $G_1, \ldots, G_k$ be subgroups of G. Prove that $G_1 + \cdots + G_k$ is a subgroup of G.

3. Let G be an abelian group, written additively, and let $G_1, \ldots, G_k$ be subgroups of G such that $G = G_1 + \cdots + G_k$. Prove that $|G| \leq |G_1| \cdots |G_k|$. Prove that $G = G_1 \oplus \cdots \oplus G_k$ if and only if $|G| = |G_1| \cdots |G_k|$.

4. Let G be an abelian group, written additively, and let $G_1, \ldots, G_k$ be subgroups of G such that $G = G_1 + \cdots + G_k$. Prove that $G = G_1 \oplus \cdots \oplus G_k$ if and only if the only representation of 0 in the form $0 = g_1 + \cdots + g_k$ with $g_i \in G_i$ is $g_1 = \cdots = g_k = 0$.

5. Let $G_1, \ldots, G_k$ be subgroups of an abelian group G such that $G = G_1 \oplus \cdots \oplus G_k$. Prove that $G \cong G_1 \times \cdots \times G_k$.

6. Let G be an additive abelian group. For every prime number p, let $G(p)$ denote the set of all elements of G whose order is a power of p. Prove that $G(p)$ is a subgroup of G.

7. Let $f : G \to H$ be a group homomorphism, and let $g \in G$. Prove that the order of $f(g)$ in H divides the order of g in G. Prove that if G is a p-group and f is surjective, then H is a p-group.

8. Let G be a finite abelian p-group. If $r_1, \ldots, r_k$ are positive integers with $r_1 \geq \cdots \geq r_k$, then we say that G is of type $(p^{r_1}, \ldots, p^{r_k})$ if $G \cong G_1 \oplus \cdots \oplus G_k$, where G_i is a cyclic group of order p^{r_i} for $i = 1, \ldots, k$. We shall prove that every finite abelian p-group has a unique type.

   Let $pG = \{pg : g \in G\}$.

   (a) Prove that pG is a subgroup of G.

   (b) Prove that if G is of type $(p^{r_1}, \ldots, p^{r_k})$ with $r_j \geq 2$ and $r_{j+1} = \cdots r_k = 1$, then pG is of type $(p^{r_1 - 1}, \ldots, p^{r_j - 1})$.

   (c) Prove that
   $$|G| = p^k |pG|.$$

   (d) Prove that if G is of type $(p^{r_1}, \ldots, p^{r_k})$ and also of type $(p^{s_1}, \ldots, p^{s_\ell})$, then $k = \ell$.

(e) Prove that if the finite abelian p-group G is of type $(p^{r_1}, \ldots, p^{r_k})$ and of type $(p^{s_1}, \ldots, p^{s_k})$, then $r_i = s_i$ for $i = 1, \ldots, k$.

Hint: Use induction on the cardinality of G. Let j and ℓ be the greatest integers such that $r_j \geq 2$ and $s_\ell \geq 2$, respectively. Apply the induction hypothesis to pG to show that $j = \ell$ and $r_i = s_i$ for $i = 1, \ldots, j$.

4.2 Characters of Finite Abelian Groups

Let G be a finite abelian group, written additively. A *group character* is a homomorphism $\chi : G \to \mathbf{C}^\times$, where $\mathbf{C}^\times$ is the multiplicative group of nonzero complex numbers. Then $\chi(0) = 1$ and $\chi(g_1 + g_2) = \chi(g_1)\chi(g_2)$ for all $g_1, g_2 \in G$.

If χ is a character of a multiplicative group G, then $\chi(1) = 1$ and $\chi(g_1 g_2) = \chi(g_1)\chi(g_2)$ for all $g_1, g_2 \in G$.

We define the character χ_0 on G by $\chi_0(g) = 1$ for all $g \in G$.

If G is an additive group of order n and if $g \in G$ has order d, then

$$\chi(g)^d = \chi(dg) = \chi(0) = 1,$$

and so $\chi(g)$ is a dth root of unity. By Theorem 2.16, d divides n and $\chi(g)$ is an nth root of unity for every $g \in G$. We have $|\chi(g)| = 1$ for all $g \in G$.

We define the product of two characters χ_1 and χ_2 by

$$\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$$

for all $g \in G$. This product is associative and commutative. The character χ_0 is a multiplicative identity, since

$$\chi_0\chi(g) = \chi_0(g)\chi(g) = \chi(g)$$

for every character χ and $g \in G$.

The inverse of the character χ is the character χ^{-1} defined by

$$\chi^{-1}(g) = \chi(-g),$$

since

$$\begin{aligned} \chi\chi^{-1}(g) &= \; = \chi(g)\chi^{-1}(g) = \chi(g)\chi(-g) \\ &= \chi(g - g) = \chi(0) = 1 \\ &= \chi_0(g), \end{aligned}$$

and so $\chi\chi^{-1} = \chi_0$.

The complex conjugate of a character χ is the character $\overline{\chi}$ defined by

$$\overline{\chi}(g) = \overline{\chi(g)}.$$

Since $|\chi(g)| = 1$ for all $g \in G$, we have

$$(\chi\overline{\chi})(g) = \chi(g)\overline{\chi})(g) = |\chi(g)|^2 = 1 = \chi_0(g),$$

and so

$$\chi^{-1}(g) = \overline{\chi}(g)$$

for every character χ and all $g \in G$.

It follows that the set of all characters of a finite abelian group G is an abelian group, called the *dual group* or *character group* of G, and denoted by $\widehat{G}$. We shall prove that $G \cong \widehat{G}$ for every finite abelian group G. We begin with finite cyclic groups.

Lemma 4.2 *The dual of a cyclic group of order n is also a cyclic group of order n.*

Proof. We introduce the exponential functions

$$e(x) = e^{2\pi ix}$$

and

$$e_n(x) = e(x/n) = e^{2\pi ix/n}.$$

The nth roots of unity are the complex numbers $e_n(a)$ for $a = 0, 1, \ldots, n-1$.

Let G be a finite cyclic group of order n with generator g_0. Then $G = \{jg_0 : j = 0, 1, \ldots, n-1\}$. For every integer a, we define $\psi_a \in \widehat{G}$ by

$$\psi_a(jg_0) = e_n(aj). \tag{4.1}$$

By Exercise 3, we have $\psi_a\psi_b = \psi_{a+b}$, $\psi_a^{-1} = \psi_{-a}$, $\psi_a = \psi_b$ if and only if $a \equiv b \pmod{n}$. It follows that

$$\psi_a = \psi_1^a$$

for every integer a. If χ is a character in $\widehat{G}$, then χ is completely determined by its value on g_0. Since $\chi(g_0)$ is an nth root of unity, we have $\chi(g_0) = e_n(a)$ for some integer $a = 0, 1, \ldots, n-1$, and so $\chi(jg_0) = e_n(aj)$ for every integer j. Therefore, $\chi = \psi_a$ and

$$\widehat{G} = \{\psi_a : a = 0, 1, \ldots, n-1\} = \{\psi_1^a : a = 0, 1, \ldots, n-1\}$$

is also a cyclic group of order n, that is, $G \cong \widehat{G}$. □

It is a simple but critical observation that if g is a nonzero element of a cyclic group G, then $\psi_1(g) \neq 1$ (Exercise 4).

Lemma 4.3 *Let G be a finite abelian group and let $G_1, \ldots, G_k$ be subgroups of G such that $G = G_1 \oplus \cdots \oplus G_k$. For every character $\chi \in \widehat{G}$ there exist unique characters $\chi_i \in \widehat{G_i}$ such that if $g \in G$ and $g = g_1 + \cdots + g_k$ with $g_i \in G_i$ for $i = 1, \ldots, k$, then*

$$\chi(g) = \chi_1(g_1) \cdots \chi_k(g_k). \tag{4.2}$$

Moreover,

$$\widehat{G} \cong \widehat{G_1} \times \cdots \times \widehat{G_k}.$$

Proof. If $\chi_i \in \widehat{G_i}$ for $i = 1, \ldots, k$, then we can construct a map $\chi : G \to \mathbf{C}^\times$ as follows. Let $g \in G$. There exist unique elements $g_i \in G_i$ such that $g = g_1 + \cdots + g_k$. Define

$$\chi(g) = \chi(g_1 + \cdots + g_k) = \chi_1(g_1) \cdots \chi_k(g_k).$$

Then χ is a character in $\widehat{G}$, and this construction induces a map

$$\Psi : \widehat{G_1} \times \cdots \times \widehat{G_k} \to \widehat{G}. \tag{4.3}$$

By Exercise 5, the map Ψ is a one-to-one homomorphism. We shall show that the map Ψ is onto. Let $\chi \in \widehat{G}$. We define the function χ_i on G_i by

$$\chi_i(g_i) = \chi(g_i) \qquad \text{for all } g_i \in G_i.$$

Then χ_i is a character in $\widehat{G_i}$. If $g \in G$ and $g = g_1 + \cdots + g_k$ with $g_i \in G_i$, then

$$\chi(g) = \chi(g_1 + \cdots + g_k) = \chi(g_1) \cdots \chi(g_k) = \chi_1(g_1) \cdots \chi_k(g_k).$$

It follows that

$$\Psi(\chi_1, \ldots, \chi_k) = \chi,$$

and so Ψ is onto. □

Theorem 4.4 *Let G be a finite abelian group. If g is a nonzero element of G, then there is a character $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$.*

Proof. We write $G = G_1 \oplus \cdots \oplus G_k$ as a direct product of cyclic groups. If $g \neq 0$, then there exist $g_1 \in G_1, \ldots, g_k \in G_k$ such that $g = g_1 + \cdots + g_k$, and $g_j \neq 0$ for some j. Since the group G_j is cyclic, there is a character $\chi_j \in \widehat{G_j}$ such that $\chi_j(g_j) \neq 1$. For $i = 1, \ldots, k, i \neq j$, let $\chi_i \in \widehat{G_i}$ be the character defined by $\chi_i(g_1) = 1$ for all $g_i \in G_i$. If $\chi = \Psi(\chi_1, \ldots, \chi_k) \in \widehat{G}$, then $\chi(g) = \chi_j(g_j) \neq 1$. □

Theorem 4.5 *A finite abelian group G is isomorphic to its dual, that is,*

$$G \cong \widehat{G}.$$

Proof. By Lemma 4.2, the dual of a finite cyclic group of order n is also a finite cyclic group of order n. By Theorem 4.3, a finite abelian group G has cyclic subgroups $G_1, \ldots, G_k$ such that

$$G = G_1 \oplus \cdots \oplus G_k.$$

By Lemma 4.3 and Exercise 5 in Section 4.1,

$$\widehat{G} \cong \widehat{G_1} \times \cdots \times \widehat{G_k} \cong G_1 \times \cdots \times G_k \cong G_1 \oplus \cdots \oplus G_k = G.$$

This completes the proof. □

Let G be a finite abelian group of order n. There is a *pairing* $\langle\ ,\ \rangle$ from $G \times \widehat{G}$ into the group of nth roots of unity defined by

$$\langle a, \chi \rangle = \chi(a).$$

This map is *nondegenerate* in the sense that $\langle a, \chi \rangle = 1$ for all group elements $a \in G$ if and only if $\chi = \chi_0$, and $\langle a, \chi \rangle = 1$ for all characters $\chi \in \widehat{G}$ if and only if $a = 0$ (by Theorem 4.4).

For each $a \in G$, the function $\langle a,\ \rangle$ is a character of the dual group $\widehat{G}$, that is, $\langle a,\ \rangle \in \widehat{\widehat{G}}$. The map $\Delta : G \to \widehat{\widehat{G}}$ defined by $a \longmapsto \langle a,\ \rangle$ or, equivalently,

$$\Delta(a)(\chi) = \langle a, \chi \rangle = \chi(a), \tag{4.4}$$

is a homomorphism of the group G into its *double dual* $\widehat{\widehat{G}}$. Since the pairing is nondegenerate, this homomorphism is one-to-one. Since $|G| = |\widehat{G}| = |\widehat{\widehat{G}}|$, it follows that Δ is a natural isomorphism of G onto $\widehat{\widehat{G}}$.

Theorem 4.6 (Orthogonality relations) *Let G be a finite abelian group of order n, and let $\widehat{G}$ be its dual group. If $\chi \in \widehat{G}$, then*

$$\sum_{a \in G} \chi(a) = \begin{cases} n & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0. \end{cases}$$

If $a \in G$, then

$$\sum_{\chi \in \widehat{G}} \chi(a) = \begin{cases} n & \text{if } a = 0, \\ 0 & \text{if } a \neq 0. \end{cases}$$

Proof. For $\chi \in \widehat{G}$, let

$$S(\chi) = \sum_{a \in G} \chi(a).$$

If $\chi = \chi_0$, then $S(\chi_0) = |G| = n$. If $\chi \neq \chi_0$, then $\chi(b) \neq 1$ for some $b \in G$, and

$$\begin{aligned} \chi(b)S(\chi) &= \chi(b) \sum_{a \in G} \chi(a) \\ &= \sum_{a \in G} \chi(ba) \\ &= \sum_{a \in G} \chi(a) \\ &= S(\chi), \end{aligned}$$

and so $S(\chi) = 0$.

For $a \in G$, let

$$T(a) = \sum_{\chi \in \widehat{G}} \chi(a).$$

If $a = 0$, then $T(a) = |\widehat{G}| = n$. If $a \neq 0$, then $\chi'(a) \neq 1$ for some $\chi' \in \widehat{G}$ (by Theorem 4.4), and

$$\begin{aligned} \chi'(a)T(a) &= \chi'(a) \sum_{\chi \in \widehat{G}} \chi(a) \\ &= \sum_{\chi \in \widehat{G}} \chi'\chi(a) \\ &= \sum_{\chi \in \widehat{G}} \chi(a) \\ &= T(a), \end{aligned}$$

and so $T(a) = 0$. This completes the proof. □

Theorem 4.7 (Orthogonality relations) *Let G be a finite abelian group of order n, and let $\widehat{G}$ be its dual group. If $\chi_1, \chi_2 \in \widehat{G}$, then*

$$\sum_{a \in G} \chi_1(a)\overline{\chi_2}(a) = \begin{cases} n & \text{if } \chi_1 = \chi_2, \\ 0 & \text{if } \chi_1 \neq \chi_2. \end{cases}$$

If $a, b \in G$, then

$$\sum_{\chi \in \widehat{G}} \chi(a)\overline{\chi}(b) = \begin{cases} n & \text{if } a = b, \\ 0 & \text{if } a \neq b. \end{cases}$$

Proof. These identities follow immediately from Theorem 4.6, since

$$\chi_1(a)\overline{\chi_2}(a) = \chi_1\chi_2^{-1}(a)$$

and

$$\chi(a)\overline{\chi}(b) = \chi(a-b).$$

This completes the proof. □

The *character table* for a group has one column for each element of the group and one row for each character of the group. For example, if C_4 is the cyclic group of order 4 with generator g_0, then the characters of C_4 are the functions

$$\psi_a(jg_0) = e_4(aj) = i^{aj}$$

for $a = 0, 1, 2, 3$, and the character table is the following.

	0	g_0	$2g_0$	$3g_0$
ψ_0	1	1	1	1
ψ_1	1	i	-1	$-i$
ψ_2	1	-1	1	-1
ψ_3	1	$-i$	-1	i

Note the that sum of the numbers in the first row is equal to the order of the group, and the sum of the numbers in each of the other rows is 0. Similarly, the sum of the numbers in the first column is the order of the group, and the sum of the numbers in each of the other columns is 0. This is a special case of the orthogonality relations.

Exercises

1. Let C_2 be the cyclic group of order 2.
   (a) Compute the character table for C_2.
   (b) Compute the character table for the group $C_2 \times C_2$.
2. Compute the character table for the cyclic group of order 6.
3. Let G be a finite cyclic group of order n. Define the characters ψ_a on G by (4.1). Prove that
   (a) $\psi_a\psi_b = \psi_{a+b}$,

(b) $\psi_a^{-1} = \psi_{-a}$,

(c) $\psi_a = \psi_b$ if and only if $a \equiv b \pmod{n}$.

4. Prove that if G is cyclic and $g \in G, g \neq 0$, then $\psi_1(g) \neq 1$.

5. Prove that the map Ψ defined by 4.3 is a one-to-one homomorphism.

6. Consider the map $\langle \quad , \quad \rangle : G \times \widehat{G} \to \mathbf{C}^\times$ defined by

$$\langle g, \chi \rangle = \chi(g).$$

Prove that

$$\langle g + g', \chi \rangle = \langle g, \chi \rangle \langle g', \chi \rangle \quad \text{and} \quad \langle g, \chi\chi' \rangle = \langle g, \chi \rangle \langle g, \chi' \rangle$$

for all $g'g' \in G$ and $\chi, \chi' \in \widehat{G}$.

7. Let $G = \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/m\mathbf{Z}$. For integers a and b, we define the function $\psi_{a,b}$ on G by

$$\psi_{a,b}(x + m\mathbf{Z}, y + m\mathbf{Z}) = e^{2\pi i(ax+by)/m} = e_m(ax + by).$$

(a) Prove that $\psi_{a,b}$ is well-defined.

(b) Prove that $\psi_{a,b} = \psi_{c,d}$ if and only if $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$.

(c) Prove that $\psi_{a,b}$ is a character of the group G.

(d) Prove that $\widehat{G} = \{\psi_{a,b} : a, b = 0, 1, \ldots, m - 1\}$.

8. Let p be a prime number, and let $G = (\mathbf{Z}/p\mathbf{Z})^\times$ be the multiplicative group of units in the field $\mathbf{Z}/p\mathbf{Z}$. Let g be a primitive root modulo p. For every integer a, define the function $\chi_a : G \to \mathbf{C}^\times$ as follows: If $(x, p) = 1$ and $x \equiv g^y \pmod{p}$, then

$$\chi_a(x + p\mathbf{Z}) = e^{2\pi ay/(p-1)} = e_{p-1}(ay).$$

(a) Prove that χ_a is a character, that is, $\chi_a \in \widehat{G}$.

(b) Prove that $\chi_a = \chi_b$ if and only if $a \equiv b \pmod{p-1}$.

(c) Prove that $\widehat{G} = \{\chi_a : a = 0, 1, \ldots, p - 2\}$.

9. Let G be a finite abelian group. For every integer r, let

$$G^r = \{rg : g \in G\}$$

and

$$\widehat{G}_r = \{\chi \in \widehat{G} : \chi^r = \chi_0\}.$$

(a) Prove that G^r is a subgroup of G and $\widehat{G}_r$ is a subgroup of $\widehat{G}$.

(b) Let $d = (r, n)$. Prove that $G^r = G^d$ and $\widehat{G}_r = \widehat{G}_d$.

(c) Let $\chi \in \widehat{G}$. Prove that $\chi \in \widehat{G}_r$ if and only if $\chi(a) = 1$ for all $a \in G^r$.

(d) Let $\chi \in \widehat{G}_r$. Define the function χ_r on the quotient group G/G^r by

$$\chi_r(a + G^r) = \chi(a).$$

Prove that χ_r is well-defined. Prove that $\chi_r \in \widehat{G/G_r}$ and that the map from $\widehat{G}_r$ to $\widehat{G/G_r}$ defined by $\chi \mapsto \chi_r$ is a group isomorphism.

10. Let G be a finite abelian group and $G^r = \{rg : g \in G\}$. Let $[G : G^r]$ be the index of the subgroup G^r in G. Prove that

$$\sum_{\chi \in \widehat{G}_r} \chi(a) = \begin{cases} [G : G^r] & \text{if } a \in G^r \\ 0 & \text{if } a \notin G^r. \end{cases}$$

Hint: Consider the quotient group G/G^r, and note that $\left|\widehat{G}_r\right| = \left|\widehat{G/G_r}\right| = [G : G^r]$.

4.3 Elementary Fourier Analysis

Let G be a finite abelian group of order n, and let $L^2(G)$ denote the n-dimensional vector space of complex-valued functions f on G. The complex conjugate of $f \in L^2(G)$ is the function $\overline{f} \in L^2(G)$ defined by

$$\overline{f}(x) = \overline{f(x)}$$

for all $x \in G$.

For $a \in G$, we define the function $\delta_a \in L^2(G)$ by

$$\delta_a(x) = \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{if } x \neq a. \end{cases}$$

If $f \in L^2(G)$, then

$$f = \sum_{a \in G} f(a)\delta_a,$$

and the set of n functions $\{\delta_a : a \in G\}$ is a basis for the vector space $L^2(G)$.

We define a function μ on the subsets of G by

$$\mu(U) = |U|$$

for all $U \subseteq G$. Then $\mu(G) = n$, and μ is *additive* in the sense that, if U_1 and U_2 are disjoint subsets of G, then $\mu(U_1 \cup U_2) = \mu(U_1) + \mu(U_2)$. The function

μ is also *translation invariant*, since $\mu(a+U) = \mu(U)$ for all $U \subseteq G$ and $a \in G$. We call μ a *Haar measure* on the group G.[1]

Using the measure μ, we define the *integral* of $f \in L^2(G)$ as

$$\int_G f = \int_G f(x)dx = \sum_{x \in G} f(x).$$

We define an inner product on the space $L^2(G)$ by

$$(f_1, f_2) = \int_G f_1 \overline{f_2} = \sum_{x \in G} f_1(x)\overline{f_2(x)}.$$

Then

$$(\delta_a, \delta_b) = \sum_{x \in G} \delta_a(x)\delta_b(x) = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{if } a \neq b, \end{cases}$$

and so the set of functions $\{\delta_a : a \in G\}$ is an orthonormal basis for $L^2(G)$. Moreover, for all $f \in L^2(G)$ and $a \in G$, we have

$$(f, \delta_a) = \sum x \in G f(x)\delta_a(x) = f(a).$$

The L^2-*norm* of a function $f \in L^2(G)$ defined by

$$\|f\|_2 = (f, f)^{1/2} = \left(\sum_{x \in G} |f(x)|^2 \right)^{1/2}.$$

The Cauchy-Schwarz inequality states that

$$|(f_1, f_2)| \leq \|f_1\|_2 \|f_2\|_2 \tag{4.5}$$

for all functions $f_1, f_2 \in L^2(G)$ (Exercise 5).

A character is a complex-valued function on G, and so $\widehat{G} \subseteq L^2(G)$. We shall show that $\widehat{G}$ is also a basis for $L^2(G)$.

If χ_1, χ_2 are characters of G, then the orthogonality relations (Theorem 4.7) imply that

$$\begin{aligned} (\chi_1, \chi_2) &= \int_G \chi_1 \overline{\chi_2} \\ &= \sum_{a \in G} \chi_1(a)\overline{\chi_2}(a) \\ &= \begin{cases} n & \text{if } \chi_1 = \chi_2 \\ 0 & \text{if } \chi_1 \neq \chi_2, \end{cases} \end{aligned}$$

[1] We can also define a measure μ on G by $\mu(U) = |U|/n$. This has the advantage that $\mu(G) = 1$, but it is not the traditional choice in elementary number theory.

and so the n characters in the dual group $\widehat{G}$ are orthogonal in the vector space $L^2(G)$. Since $|\widehat{G}| = |G| = \dim_{\mathbf{C}} L^2(G) = n$, it follows that $\widehat{G}$ is a basis for $L^2(G)$.

There are an analogous Haar measure and inner product on the dual group $\widehat{G}$. If $\widehat{f}, \widehat{f}_2 \in L^2(\widehat{G})$, then

$$(\widehat{f}_1, \widehat{f}_2) = \int_{\widehat{G}} \widehat{f}_1 \overline{\widehat{f}_2} = \sum_{\chi \in \widehat{G}} \widehat{f}_1(\chi) \overline{\widehat{f}_2}(\chi).$$

Let $\widehat{\widehat{G}}$ denote the double dual of G, that is, the group of characters of the dual group $\widehat{G}$. For $a \in G$, we defined $\Delta(a) \in \widehat{\widehat{G}}$ by

$$\Delta(a)(\chi) = \chi(a),$$

and we proved that every character in $\widehat{\widehat{G}}$ is of the form $\Delta(a)$ for some $a \in G$. By the orthogonality relations (Theorem 4.7), for every $a, b \in G$ we have

$$\begin{aligned} (\Delta(a), \Delta(b))_{\widehat{G}} &= \sum_{\chi \in \widehat{G}} \Delta(a)(\chi) \overline{\Delta(b)}(\chi) \\ &= \sum_{\chi \in \widehat{G}} \chi(a) \overline{\chi}(b) \\ &= \begin{cases} n & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases} \end{aligned}$$

The *Fourier transform* is a linear transformation from $L^2(G)$ to $L^2(\widehat{G})$ that sends the function $f \in L^2(G)$ to the function $\widehat{f} \in L^2(\widehat{G})$, where

$$\widehat{f}(\chi) = (f, \chi) = \sum_{g \in G} f(g) \overline{\chi}(g). \tag{4.6}$$

For example, the Fourier transform of the function $\delta_a \in L^2(G)$ is

$$\widehat{\delta_a}(\chi) = \sum_{g \in G} \delta_a(g) \overline{\chi}(g) = \overline{\chi}(a) = \chi(-a).$$

The process of recovering f from its Fourier transform $\widehat{f}$ is called *Fourier inversion.*

Theorem 4.8 (Fourier inversion) *Let G be a finite abelian group of order n with dual group $\widehat{G}$. If $f \in L^2(G)$, then*

$$f = \frac{1}{n} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi) \chi, \tag{4.7}$$

and (4.7) is the unique representation of f as a linear combination of characters of G.

Let $\Delta : G \to \widehat{\widehat{G}}$ be the isomorphism defined by $\Delta(a)(\chi) = \chi(a)$ for all $\chi \in \widehat{G}$. If $f \in L^2(G)$, then $\widehat{\widehat{f}} \in L^2\left(\widehat{\widehat{G}}\right)$, and, for every $a \in G$,

$$\widehat{\widehat{f}}(\Delta(a)) = nf(-a). \tag{4.8}$$

Proof. This is a straightforward calculation. Let $a \in G$. Defining the Fourier transform by (4.6), we have

$$\begin{aligned}\frac{1}{n}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(a) &= \frac{1}{n}\sum_{\chi\in\widehat{G}}\left(\sum_{b\in G} f(b)\overline{\chi}(b)\right)\chi(a) \\ &= \sum_{b\in G} f(b)\left(\frac{1}{n}\sum_{\chi\in\widehat{G}}\chi(a)\overline{\chi}(b)\right) \\ &= f(a),\end{aligned}$$

by the orthogonality relations (Theorem 4.7). This proves (4.7). The uniqueness of the series (4.7) is Exercise 2.

To prove (4.8), we have

$$\begin{aligned}\widehat{\widehat{f}}(\Delta(a)) &= \sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\overline{\Delta(a)}(\chi) \\ &= \sum_{\chi\in\widehat{G}}\sum_{g\in G} f(g)\overline{\chi}(g)\overline{\chi}(a) \\ &= \sum_{g\in G} f(g)\sum_{\chi\in\widehat{G}}\overline{\chi}(g+a) \\ &= nf(-a).\end{aligned}$$

This completes the proof. □

The sum (4.7) is called the *Fourier series* for the function f.

Theorem 4.9 (Plancherel's formula) *If G is a finite abelian group of order n and $f \in L^2(G)$, then*

$$\|\widehat{f}\|_2 = \sqrt{n}\|f\|_2.$$

Proof. We have

$$\|\widehat{f}\|_2^2 = (\widehat{f}, \widehat{f})$$

$$
\begin{aligned}
&= \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\overline{\widehat{f}(\chi)} \\
&= \sum_{\chi \in \widehat{G}} \left(\sum_{b \in G} f(b)\overline{\chi(b)} \right) \left(\sum_{a \in G} \overline{f(a)}\chi(a) \right) \\
&= \sum_{a \in G} \sum_{b \in G} \overline{f(a)} f(b) \left(\sum_{\chi \in \widehat{G}} \chi(a)\overline{\chi(b)} \right) \\
&= n \sum_{a \in G} |f(a)|^2 \\
&= n\|f\|_2^2.
\end{aligned}
$$

This completes the proof. □

Let G be a finite abelian group of order $|G| = n$, and let $f \in L^2(G)$. The *support* of f is the set

$$\text{supp}(f) = \{a \in G : f(a) \neq 0\}.$$

We define the L^∞-*norm* of a function $f \in L^2(G)$ by

$$\|f\|_\infty = \max\{|f(a)| : a \in G\}.$$

For every function $f \in L^2(G)$ we have the elementary inequality

$$\|f\|_2^2 = (f, f) = \sum_{a \in G} |f(a)|^2 \leq \|f\|_\infty^2 |\text{supp}(f)|. \tag{4.9}$$

The *uncertainty principle* in Fourier analysis states that if $f \in L^2(G)$ is a function with Fourier transform $\widehat{f} \in L^2(\widehat{G})$, then the sets $\text{supp}(f)$ and $\text{supp}(\widehat{f})$ cannot be simultaneously small. This has the following quantitative formulation.

Theorem 4.10 (Uncertainty principle) *If G is a finite abelian group and $f \in L^2(G)$, $f \neq 0$, then*

$$|supp(f)||supp(\widehat{f})| \geq |G|.$$

Proof. Let $a \in G$. By Theorem 4.8,

$$f(a) = \frac{1}{n} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(a).$$

Since $|\chi(a)| = 1$ for all $\chi \in \widehat{G}$, it follows that

$$|f(a)| \leq \frac{1}{n} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)| = \frac{1}{n} \sum_{\chi \in \text{supp}(\widehat{f})} |\widehat{f}(\chi)|$$

and so

$$\|f\|_\infty \leq \frac{1}{n} \sum_{\chi \in \mathrm{supp}(\widehat{f})} |\widehat{f}(\chi)|.$$

Applying the Cauchy-Schwarz inequality (4.5) with $f_1 = \widehat{f}(\chi)$ and with f_2 the characteristic function of the set supp $\left(\widehat{f}\right)$, we have

$$\left(\sum_{\chi \in \mathrm{supp}(\widehat{f})} |\widehat{f}(\chi)| \right)^2 = \sum_{\chi \in \mathrm{supp}(\widehat{f})} |\widehat{f}(\chi)|^2 |\mathrm{supp}(\widehat{f})|.$$

Using Plancherel's formula (Theorem 4.9), and inequality (4.9), we obtain

$$\begin{aligned}
\|f\|_\infty^2 &\leq \frac{1}{n^2} \left(\sum_{\chi \in \mathrm{supp}(\widehat{f})} |\widehat{f}(\chi)| \right)^2 \\
&\leq \frac{1}{n^2} \sum_{\chi \in \mathrm{supp}(\widehat{f})} |\widehat{f}(\chi)|^2 |\mathrm{supp}(\widehat{f})| \\
&= \frac{1}{n^2} \|\widehat{f}\|_2^2 |\mathrm{supp}(\widehat{f})| \\
&= \frac{1}{n} \|f\|_2^2 |\mathrm{supp}(\widehat{f})| \\
&\leq \frac{1}{n} \|f\|_\infty^2 |\mathrm{supp}(f)| |\mathrm{supp}(\widehat{f})|.
\end{aligned}$$

Since $f \neq 0$, we have $\|f\|_\infty > 0$ and so

$$|\mathrm{supp}(f)||\mathrm{supp}(\widehat{f})| \geq n = |G|.$$

This completes the proof. □

If $f \in L^2(G)$ and $|\mathrm{supp}(f)| = 1$, then the uncertainty principle implies that $|supp(\widehat{f})| = |G|$, that is, $\widehat{f}(\chi) \neq 0$ for all $\chi \in \widehat{G}$. Here is an example. Let $a \in G$ and $f = \delta_a \in L^2(G)$. Then $\delta_a(x) \neq 0$ if and only if $x = a$, and so $|\mathrm{supp}(\delta_a)| = 1$. We have $\widehat{\delta_a}(\chi) = \overline{\chi}(a) \neq 0$ for all $\chi \in \widehat{G}$. This shows that the lower bound in the uncertainty principle is best possible.

Exercises

In these exercises, G is a finite abelian group.

1. Let $f, g \in L^2(G)$. Prove that

$$(g, f) = \overline{(f, g)}.$$

2. Let $f \in L^2(G)$. Prove that if $c \in L^2(\widehat{G})$ and $f = (1/n)\sum_{\chi \in \widehat{G}} c(\chi)\chi$, then $c(\chi) = \widehat{f}(\chi)$.

3. Prove that the Haar measure on G is unique, that is, there exists a unique function μ on the subsets of G such that μ is additive, translation invariant, and $\mu(G) = n$.

4. Let $U : L^2(G) \to L^2(\widehat{G})$ be a linear transformation such that $U(\delta_a)(\chi) = \overline{\chi}(a)$ for all $\chi \in \widehat{G}$. Prove that U is the Fourier transform, that is, $U(f) = \widehat{f}$ for all $f \in L^2(G)$.

5. (Cauchy-Schwarz inequality) Let $f, g \in L^2G$. Prove that

$$|(f, g)| \leq \|f\|_2 \|g\|_2.$$

Hint: If $\lambda \in \mathbf{C}$, then $\|f - \lambda g\|_2^2 \geq 0$. For $g \neq 0$, apply this inequality with $\lambda = (f, g)/(g, g)$.

6. Prove that if $f, g \in L^2(G)$, then

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

7. Let $\chi_1, \chi_2 \in \widehat{G}$. Prove that

$$\widehat{\chi_1}(\chi_2) = \begin{cases} n & \text{if } \chi_1 = \chi_2 \\ 0 & \text{if } \chi_1 \neq \chi_2. \end{cases}$$

8. Use the uncertainty principle to prove that the Fourier transform is one-to-one.

Hint: Prove that if $f \in L^2(G)$ and $f \neq 0$, then $\widehat{f} \neq 0$.

9. For $a \in G$ and $f \in L^2(G)$, we define the translation operator T_a on $L^2(G)$ by $T_a(f)(x) = f(x - a)$. Prove that $\widehat{T_a(f)} = \overline{\chi}(a)\widehat{f}$.

10. For functions $f_1, f_2 \in L^2(G)$, we define the *convolution* $f_1 * f_2 \in L^2(G)$ by

$$f_1 * f_2(a) = \int_G f_1(a - x) f_2(x) dx = \sum_{a \in G} f_1(a - x) f_2(x).$$

(a) Prove that

$$f_1 * f_2(a) = \sum_{x+y=a} f_1(x) f_2(y).$$

(b) Prove that convolution is commutative, that is,

$$f_1 * f_2 = f_2 * f_1.$$

(c) Prove that convolution is associative, that is,

$$(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3).$$

(d) Prove that, if $f_1, \ldots, f_k \in L^2(G)$, then

$$f_1 * \cdots * f_k(a) = \sum_{x_1+\cdots+x_k=a} f_1(x_1)\cdots f_k(x_k).$$

11. Let $\chi \in \widehat{G}$. Prove that

$$\underbrace{\chi * \cdots * \chi}_{k \text{ times}}(a) = \sum_{x_1+x_2+\cdots+x_k=a} \chi(x_1 + x_2 + \cdots + x_k).$$

12. Let p be a prime number, and define $\ell_p \in L^2(\mathbf{Z}/p\mathbf{Z})$ by

$$\ell_p(a + p\mathbf{Z}) = \left(\frac{a}{p}\right),$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. Prove that

$$\underbrace{\ell_p * \cdots * \ell_p}_{k \text{ times}}(a + p\mathbf{Z}) = \sum_{\substack{x_1+x_2+\cdots+x_k=a \\ 1\le x_i\le p-1}} \left(\frac{x_1x_2\cdots x_k}{p}\right).$$

13. Let $f_1, f_2, \ldots, f_k \in L^2(G)$. Prove that a product of Fourier transforms is the convolution of the product in the sense that

$$\widehat{f_1} \cdot \widehat{f_2} = \widehat{f_1 * f_2}$$

and

$$\widehat{f_1} \cdot \widehat{f_2} \cdots \widehat{f_k} = \widehat{f_1 * f_2 * \cdots * f_k}.$$

14. Prove that $\delta_a * f = T_a(f)$ for all $f \in L^2(G)$. Use this to give another proof of Exercise 9.

4.4 Poisson Summation

Let G be a finite abelian group with subgroup H, and let $L^2(G)^H$ be the vector space of complex-valued functions on G that are constant on cosets in G/H, that is,

$$L^2(G)^H = \{f \in L^2(G) : f(x + h) = f(x) \text{ for all } x \in G \text{ and } h \in H\}.$$

Let $\widehat{G}^H$ be the group of characters of G that are trivial on H, that is,

$$\widehat{G}^H = \{\chi \in \widehat{G} : \chi(h) = 1 \text{ for all } h \in H\}.$$

Lemma 4.4 *Let G be a finite abelian group with subgroup H. Then*

$$\widehat{G}^H = \widehat{G} \cap L^2(G)^H.$$

Proof. If $\chi \in \widehat{G}^H \subseteq \widehat{G}$, then $\chi(x+h) = \chi(x)\chi(h) = \chi(x)$ for all $x \in G$ and $h \in H$, and so $\chi \in \widehat{G} \cap L^2(G/H)$. Conversely, if $\chi \in \widehat{G} \cap L^2(G/H)$, then $\chi(h) = \chi(0+h) = \chi(0) = 1$ for all $h \in H$, and $\chi \in \widehat{G}^H$. □

Lemma 4.5 *Let G be a finite abelian group with subgroup H, and let $\pi : G \to G/H$ be the natural map onto the quotient group. For $f^\sharp \in L^2(G/H)$, define the map $\pi^\sharp(f^\sharp) \in L^2(G)$ by*

$$\pi^\sharp(f^\sharp)(x) = f^\sharp\pi(x) = f^\sharp(x+H)$$

for all $x \in G$. Then $\pi^\sharp$ is a vector space isomorphism from $L^2(G/H)$ onto $L^2(G)^H$. Moreover,

$$\pi^\sharp\left(\widehat{G/H}\right) \subseteq \widehat{G}^H,$$

and the map

$$\pi^\sharp : \widehat{G/H} \to \widehat{G}^H$$

is a group isomorphism.

Proof. Let $f^\sharp \in L^2(G/H)$. If $x \in G$ and $h \in H$, then

$$\pi^\sharp(f^\sharp)(x+h) = f^\sharp\pi(x+h) = f^\sharp\pi(x) = \pi^\sharp(f^\sharp)(x),$$

and so $\pi^\sharp$ maps $L^2(G/H)$ into $L^2(G)^H$. It is easy to check that $\pi^\sharp$ is linear. Moreover, $\pi^\sharp$ is onto, since if $f \in L^2(G)^H$, then there is a well-defined map $f^\sharp \in L^2(G/H)$ given by $f^\sharp(x+H) = f(x)$, and $\pi^\sharp(f^\sharp)(x) = f^\sharp(x+H) = f(x)$ for all $x \in G$. Finally, $\pi^\sharp$ is one-to-one since $\pi^\sharp(f^\sharp)(x) = 0$ for all $x \in G$ if and only if $f^\sharp(x+H) = 0$ for all $x+H \in G/H$, that is, if and only if $f^\sharp = 0$. This proves that $\pi^\sharp$ is an isomorphism.

If $\chi^\sharp \in \widehat{G/H}$, then

$$\begin{aligned}
\pi^\sharp(\chi^\sharp)(x+y) &= \chi^\sharp(\pi(x+y)) \\
&= \chi^\sharp(x+y+H) \\
&= \chi^\sharp(x+H)\chi^\sharp(y+H) \\
&= \pi^\sharp(\chi^\sharp)(x)\pi^\sharp(\chi^\sharp)(y),
\end{aligned}$$

and so

$$\pi^\sharp(\chi^\sharp) \in \widehat{G} \cap L^2(G)^H = \widehat{G}^H.$$

It is left as an exercise to prove that $\pi^\sharp : \widehat{G/H} \to \widehat{G}^H$ is a group isomorphism (Exercise 2). □

Theorem 4.11 (Poisson summation formula) *Let G be a finite abelian group and H a subgroup of G. If $f \in L^2(G)$, then*

$$\frac{1}{|H|}\sum_{y\in H} f(y) = \frac{1}{|G|}\sum_{\chi\in \widehat{G}^H} \widehat{f}(\chi).$$

Proof. Let $f \in L^2(G)$ and $\chi \in \widehat{G}^H$. We define the function $f^\sharp \in L^2(G/H)$ by

$$f^\sharp(x+H) = \sum_{y\in H} f(x+y).$$

We define the character $\chi^\sharp \in \widehat{G/H}$ by $\chi^\sharp(x+H) = \chi(x)$. If $\pi^\sharp : \widehat{G/H} \to \widehat{G}^H$ is the isomorphism constructed in Lemma 4.5, then $\pi^\sharp(\chi^\sharp) = \chi$, and the Fourier transform of $f^\sharp$ is

$$\begin{aligned}
\widehat{f^\sharp}(\chi^\sharp) &= \sum_{x+H\in G/H} f^\sharp(x+H)\overline{\chi^\sharp}(x+H) \\
&= \sum_{x+H\in G/H}\sum_{y\in H} f(x+y)\overline{\chi}(x) \\
&= \sum_{x+H\in G/H}\sum_{y\in H} f(x+y)\overline{\chi}(x+y) \\
&= \sum_{x\in G} f(x)\overline{\chi}(x) \\
&= \widehat{f}(\chi).
\end{aligned}$$

It follows that the Fourier series for $f^\sharp$ is

$$\begin{aligned}
f^\sharp(x+H) &= \frac{1}{|G/H|}\sum_{\chi^\sharp\in\widehat{G/H}} \widehat{f^\sharp}(\chi^\sharp)\overline{\chi^\sharp}(x+H) \\
&= \frac{|H|}{|G|}\sum_{\chi\in\widehat{G}^H} \widehat{f}(\chi)\overline{\chi}(x).
\end{aligned}$$

Equivalently, for $x \in G$,

$$\frac{1}{|H|}\sum_{y\in H} f(x+y) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}^H} \widehat{f}(\chi)\overline{\chi}(x).$$

This is the Poisson summation formula. □

Exercises

In these exercises, G is a finite abelian group and H is a subgroup of G.

1. Let $\widehat{G}^H$ denote the set of all characters χ of G such that $\chi(h) = 1$ for all $h \in H$. Prove that $\widehat{G}^H$ is a subgroup of $\widehat{G}$.

2. Let $\pi^\sharp : \widehat{G/H} \to \widehat{G}^H$ be the map constructed in Lemma 4.5. Prove that $\pi^\sharp$ is a group homomorphism. Define $\lambda : \widehat{G}^H \to \widehat{G/H}$ by $\lambda(\chi)(x+H) = \chi(x)$. Prove that λ is a well-defined group homomorphism, and that $\lambda^{-1} = \pi^\sharp$.

3. Prove that G contains a subgroup isomorphic to G/H.

   Hint:
$$G/H \cong \widehat{G/H} \cong \widehat{G}^H \subseteq \widehat{G} \cong G.$$

4. To each character $\chi \in \widehat{G}$ there is a corresponding character $\chi' \in \widehat{H}$ defined by restriction:
$$\chi'(h) = \chi(h) \qquad \text{for } h \in H.$$
   Prove that this defines a homomorphism $\rho : \widehat{G} \to \widehat{H}$ with kernel $\widehat{G}^H$. This induces a one-to-one homomorphism of $\tilde{\rho} : \widehat{G}/\widehat{G}^H \to \widehat{H}$. Prove that $\tilde{\rho}$ is surjective, and so
$$\widehat{G}/\widehat{G}^H \cong \widehat{H}.$$
   Hint: These two groups have the same cardinality.

5. Let $f \in L^2(G)$, and define $f^\sharp \in L^2(G)$ by
$$f^\sharp(x) = \sum_{h \in H} f(x+h).$$
   Prove that $f^\sharp \in L^2(G)^H$ and
$$\int_G f = \frac{1}{|H|} \int_G f^\sharp.$$

6. Let G_1 and G_2 be finite abelian groups. Let $f \in L^2(G_1 \times G_2)$. For $x_1 \in G_1$, define the function $f_{x_1} \in L^2(G_2)$ by $f_{x_1}(x_2) = f(x_1, x_2)$. Show that Poisson summation applied to the group $G = G_1 \times G_2$ and subgroup $H = G_1 \times \{0\}$ gives
$$\sum_{x_1 \in G_1} f_{x_1}(0) = \frac{1}{|G_2|} \sum_{x_1 \in G_1} \sum_{\chi_2 \in \widehat{G_2}} \widehat{f_{x_1}}(\chi_2).$$

7. Let $f \in L^2(G \times G)$. Use Poisson summation to prove that

$$\sum_{x \in G} f(x,x) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \sum_{(x,y) \in G \times G} f(x,y)\chi(x)\overline{\chi}(y).$$

Note that this identity is also an immediate consequence of the orthogonality relations.

8. This is another example that shows that the lower bound in the uncertainty principle (Theorem 4.10) is best possible. Let H be a subgroup of G, and define $\delta_H \in L^2(G)$ by

$$\delta_H(x) = \begin{cases} 1 & \text{if } x \in H \\ 0 & \text{if } x \notin H. \end{cases}$$

(a) Prove that

$$\text{supp}(\delta_H) = H.$$

(b) Prove that if $\chi \in \widehat{G}$, then

$$\widehat{\delta_H}(\chi) = \begin{cases} |H| & \text{if } \chi \in \widehat{G}^H \\ 0 & \text{if } \chi \notin \widehat{G}^H. \end{cases}$$

(c) Prove that

$$\text{supp}(\delta_H)\text{supp}\left(\widehat{\delta_H}\right) = |G|.$$

4.5 Trace Formulae on Finite Abelian Groups

We recall some facts from linear algebra. Let $A = (a_{ij})$ be an $n \times n$ matrix. The *trace* of A is the sum of the diagonal elements of A, that is,

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Let $B = (b_{ij})$ be another $n \times n$ matrix. The simplest *trace formula* (Exercise 1) states that

$$\text{tr}(AB) = \text{tr}(BA). \tag{4.10}$$

Every result in this section follows from this fundamental identity.

Let V be an n-dimensional vector space, and let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for V. If $T : V \to V$ is a linear operator, and

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i,$$

then the $n \times n$ matrix $A = (a_{ij}) = [T]_{\mathcal{B}}$ is called the matrix of the operator T with respect to the basis $\mathcal{B}$.

Let $\mathcal{B}' = \{v_1', \ldots, v_n'\}$ be another basis for V, and let

$$T(v_j') = \sum_{i=1}^{n} a_{ij}' v_i'. \tag{4.11}$$

Then $A' = (a_{ij}') = [T]_{\mathcal{B}'}$ is the matrix of T with respect to the basis $\mathcal{B}$.

Each vector $v_j' \in \mathcal{B}'$ is a linear combination of the vectors in the basis $\mathcal{B}$,

$$v_j' = \sum_{i=1}^{n} r_{ij} v_i, \tag{4.12}$$

and each vector $v_j \in \mathcal{B}$ is a linear combination of the vectors in the basis $\mathcal{B}'$,

$$v_j = \sum_{i=1}^{n} s_{ij} v_i'. \tag{4.13}$$

Consider the $n \times n$ matrices $R = (r_{ij})$ and $S = (s_{ij})$. Then $S = R^{-1}$ (Exercise 2). We have

$$\begin{aligned}
T(v_j') &= T\left(\sum_{\ell=1}^{n} r_{\ell j} v_\ell\right) \\
&= \sum_{\ell=1}^{n} r_{\ell j} T(v_\ell) \\
&= \sum_{\ell=1}^{n} r_{\ell j} \sum_{k=1}^{n} a_{k\ell} v_k \\
&= \sum_{\ell=1}^{n} r_{\ell j} \sum_{k=1}^{n} a_{k\ell} \sum_{i=1}^{n} s_{ik} v_i' \\
&= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \sum_{\ell=1}^{n} s_{ik} a_{k\ell} r_{\ell j}\right) v_i'.
\end{aligned}$$

Comparing this with (4.11), we obtain

$$a_{ij}' = \sum_{k=1}^{n} \sum_{\ell=1}^{n} s_{ik} a_{k\ell} r_{\ell j}$$

for all $i, j = 1, \ldots, n$, and so

$$A' = SAR = R^{-1}AR.$$

Identity (4.10) implies that

$$tr(A') = tr(R^{-1}AR) = tr(ARR^{-1}) = \mathrm{tr}(A).$$

It follows that we can define the trace of a linear operator T on a vector space V as the trace of the matrix of T with respect to some basis for V, and that this definition does not depend on the choice of basis.

The vector $v' \in V$ is called an *eigenvector* for the operator T with *eigenvalue* λ if $v' \neq 0$ and $T(v') = \lambda v'$. The operator T is *diagonalizable* if there exists a basis for V consisting of eigenvectors, that is, there exist nonzero vectors $v'_1, \ldots, v'_n \in V$ and numbers $\lambda_1, \ldots, \lambda_n$ such that $\mathcal{B}' = \{v'_1, \ldots, v'_n\}$ is a basis for V and $T(v'_i) = \lambda_i v'_i$ for $i = 1, \ldots, n$. In this case, the matrix for T with respect to the basis $\mathcal{B}'$ is the diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix},$$

and so

$$\sum_{i=1}^{n} a_{ii} = \operatorname{tr}(A) = \operatorname{tr}(D) = \sum_{i=1}^{n} \lambda_i.$$

We restate this important identity as a theorem.

Theorem 4.12 (Elementary trace formula) *Let T be a linear operator on an n-dimensional vector space V, let $\mathcal{B}$ be a basis for V, and let $A = (a_{ij})$ be the matrix of T with respect to $\mathcal{B}$. If T is diagonalizable, then V has a basis $\mathcal{B} = \{v'_1, \ldots, v'_n\}$ of eigenvectors with $T(v_i) = \lambda_i v_i$ for $i = 1, \ldots, n$, and the trace of A is equal to the sum of the eigenvalues of T, that is,*

$$\sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i.$$

We shall show that both the Fourier inversion theorem and the Poisson summation formula are consequences of this elementary trace formula.

Let G be a finite abelian group of order n, and let $L^2(G)$ be the n-dimensional vector space of complex-valued functions on G. For every $a \in G$ there is a linear operator T_a on $L^2(G)$ defined by $T_a(f)(x) = f(x-a)$. The operator T_a is called *translation by* a.

Another class of operators on $L^2(G)$ are *integral operators*. A function $K \in L^2(G \times G)$ induces a linear operator Φ_K on the vector space $L^2(G)$ as follows: For $f \in L^2(G)$, let

$$\Phi_K(f)(x) = \int_G K(x,y) f(y) dy = \sum_{y \in G} K(x,y) f(y).$$

The map Φ_K is called an *integral operator* on $L^2(G)$ with *kernel* $K(x,y)$.

Let $G = \{x_1, \ldots, x_n\}$. Associated to the kernel K is a matrix $A = (a_{ij}) \in M_n(\mathbf{C})$ defined by

$$a_{ij} = K(x_i, x_j). \tag{4.14}$$

Conversely, to every matrix $A = (a_{ij}) \in M_n(\mathbf{C})$ there is a function $K(x, y) \in L^2(G \times G)$ defined by (4.14), and an associated integral operator Φ_K.

Theorem 4.13 *Let $G = \{x_1, \ldots, x_n\}$ be an abelian group of order n. Let $K \in L^2(G \times G)$ and let Φ_K be the associated integral operator on $L^2(G)$. The matrix of Φ_K with respect to the orthonormal basis $\{\delta_{x_i} : i = 1, \ldots, n\}$ is $(K(x_i, x_j))$, and the trace of Φ_K is*

$$tr(\Phi_K) = \sum_{i=1}^{n} K(x_i, x_i). \tag{4.15}$$

Proof. The matrix of the operator Φ_K is (c_{ij}), where c_{ij} is defined by

$$\Phi_K(\delta_{x_j}) = \sum_{i=1}^{n} c_{ij}\delta_{x_i}.$$

Then

$$c_{ij} = \Phi_K(\delta_{x_j})(x_i) = \sum_{y \in G} K(x_i, y)\delta_{x_j}(y) = K(x_i, x_j).$$

This completes the proof. □

Theorem 4.14 *Let G be a finite abelian group. Let $K \in L^2(G \times G)$ with Φ_K the associated integral operator on $L^2(G)$. The operator Φ_K commutes with all translations T_a, that is,*

$$T_a\Phi_K(f) = \Phi_K T_a(f)$$

for all $a \in G$ and $f \in L^2(G)$, if and only if there exists a function $h \in L^2(G)$ such that $K(x, y) = h(x-y)$ for all $x, y \in G$. In this case, Φ_K is convolution by h, that is,

$$\Phi_K(f)(x) = h * f(x) = \int_G h(x - y)f(y)dy,$$

and the trace of Φ_K is

$$tr(\Phi_K) = nh(0).$$

Proof. Let $f, h \in L^2(G)$. We define the convolution operator C_h on $L^2(G)$ by

$$C_h(f)(x) = h * f(x) = \int_G h(x - y)f(y)dy = \sum_{y \in G} h(x - y)f(y).$$

(See Exercise 10 in Section 4.3.) Define $K(x,y) \in L^2(G \times G)$ by $K(x,y) = h(x-y)$. Then

$$\Phi_K(f)(x) = \int_G K(x,y)f(y)dy = \int_G h(x-y)f(y)dy = C_h(f)(x),$$

and Φ_K is convolution by h. For $a, x \in G$, we have

$$\begin{aligned} T_a C_h(f)(x) &= C_h(f)(x-a) \\ &= \sum_{y \in G} h(x-a-y)f(y) \\ &= \sum_{y \in G} h(x-y)f(y-a) \\ &= \sum_{y \in G} h(x-y)T_a(f)(y) \\ &= C_h T_a(f)(x), \end{aligned}$$

and so $T_a C_h = C_h T_a$, that is, convolution commutes with translations.

Conversely, let $K(x,y) \in L^2(G \times G)$. For $a, x \in G$ and $f \in L^2(G)$, we have

$$T_a \Phi_K(f)(x) = \Phi_K(f)(x-a) = \sum_{y \in G} K(x-a,y)f(y)$$

and

$$\begin{aligned} \Phi_K T_a(f)(x) &= \sum_{y \in G} K(x,y)T_a(f)(y) \\ &= \sum_{y \in G} K(x,y)f(y-a) \\ &= \sum_{y \in G} K(x,a+y)f(y). \end{aligned}$$

If Φ_K commutes with translations, then $T_a \Phi_K = \Phi_K T_a$, and

$$\sum_{y \in G} K(x-a,y)f(y) = \sum_{y \in G} K(x,a+y)f(y).$$

Applying this identity to the function

$$f(x) = \delta_0(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0. \end{cases}$$

we obtain $K(x-a,0) = K(x,a)$ for all $a, x \in G$. Define the function $h \in L^2(G)$ by

$$h(x) = K(x,0).$$

Then

$$K(x,y) = K(x-y,0) = h(x-y)$$

for all $x, y \in G$, and the operator Φ_K is convolution by $h(x)$. Moreover, $\mathrm{tr}(\Phi_K) = nh(0)$ by (4.15). This completes the proof. □

Theorem 4.15 (Trace formula) *For $h \in L^2(G)$, let C_h be the convolution operator on $L^2(G)$, that is, $C_h(f) = h * f$ for $f \in L^2(G)$. The dual group $\widehat{G}$ is a basis of eigenvectors for C_h. If χ is a character in $\widehat{G}$, then χ has eigenvalue $\widehat{h}(\chi)$, that is,*

$$C_h(\chi) = \widehat{h}(\chi)\chi,$$

and

$$nh(0) = \sum_{\chi \in \widehat{G}} \widehat{h}(\chi).$$

Proof. This is a straightforward calculation. For $x \in G$, we have

$$\begin{aligned} C_h(\chi)(x) &= h * \chi(x) = \chi * h(x) \\ &= \sum_{y \in G} \chi(x-y)h(y) \\ &= \left(\sum_{y \in G} h(y)\overline{\chi}(y) \right) \chi(x) \\ &= \widehat{h}(\chi)\chi(x), \end{aligned}$$

and so χ is an eigenvector of the convolution C_h with eigenvalue $\widehat{h}(\chi)$. By Theorem 4.12, since $\widehat{G}$ is a basis for $L^2(G)$, the trace of C_h is the sum of the eigenvalues, that is,

$$\mathrm{tr}(C_h) = \sum_{\chi \in \widehat{G}} \widehat{h}(\chi).$$

By Theorem 4.14, we also have

$$\mathrm{tr}(C_h) = nh(0).$$

This completes the proof. □

We can immediately deduce the Fourier inversion formula (Theorem 4.8) from Theorem 4.15. If $f \in L^2(G)$, then

$$f(0) = \frac{1}{n} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi). \tag{4.16}$$

This trace formula can also be obtained by computing the Fourier series for f at $x = 0$. On the other hand, if we simply apply (4.16) to the function $T_{-a}(f)$ and use Exercise 9 in Section 4.3, then we obtain

$$\begin{aligned} f(a) &= T_{-a}(f)(0) \\ &= \frac{1}{n}\sum_{\chi\in\widehat{G}} \widehat{T_{-a}(f)}(\chi) \\ &= \frac{1}{n}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(a). \end{aligned}$$

This is the Fourier inversion formula.

Next, we derive the Poisson summation formula (Theorem 4.11) from the elementary trace formula.

Let H be a subgroup of G, and let $\pi : G \to G/H$ be the natural map. For $x \in G$, define $x^\sharp = \pi(x) = x + H \in G/H$. There is an orthonormal basis for the vector space $L^2(G/H)$ that consists of the functions $\delta_{x^\sharp}$, where

$$\delta_{x^\sharp}(y^\sharp) = \begin{cases} 1 & \text{if } x^\sharp = y^\sharp \\ 0 & \text{if } x^\sharp \neq y^\sharp. \end{cases}$$

For $f \in L^2(G)$, define the function $f^\sharp \in L^2(G/H)$ by

$$f^\sharp(x+H) = \sum_{y\in H} f(x+y).$$

Let $C_{f^\sharp}$ be convolution by $f^\sharp$ on $L^2(G/H)$. The operator $C_{f^\sharp}$ has matrix $\left(f^\sharp(x^\sharp - y^\sharp)\right)$, with respect to the basis $\{\delta_{x^\sharp}\}$. By Theorem 4.14, the trace of $C_{f^\sharp}$ is

$$\operatorname{tr}(C_{f^\sharp}) = |G/H| f^\sharp(0^\sharp) = \frac{|G|}{|H|}\sum_{y\in H} f(y).$$

By Theorem 4.15, the character group $\widehat{G/H}$ is a basis of eigenvectors for the convolution operator $C_{y^\sharp}$. If $\chi^\sharp \in \widehat{G/H}$ and $\chi = \pi^\sharp(\chi^\sharp) \in \widehat{G}^H$, then

$$C_{f^\sharp}(\chi^\sharp) = \widehat{f^\sharp}(\chi^\sharp)\chi^\sharp,$$

with eigenvalue

$$\begin{aligned} \widehat{f^\sharp}(\chi^\sharp) &= \sum_{x^\sharp\in G/H} f^\sharp(x^\sharp)\overline{\chi^\sharp}(x^\sharp) \\ &= \sum_{x^\sharp\in G/H}\sum_{y\in H} f(x+y)\overline{\chi}(x) \\ &= \sum_{x^\sharp\in G/H}\sum_{y\in H} f(x+y)\overline{\chi}(x+y) \\ &= \sum_{x\in G} f(x)\overline{\chi}(x). \end{aligned}$$

It follows that

$$\operatorname{tr}(C_{f^\sharp}) = \sum_{\chi^\sharp \in \widehat{G/H}} \widehat{f^\sharp}(\chi^\sharp) = \sum_{\chi \in \widehat{G}^H} \sum_{x \in G} f(x)\overline{\chi}(x), = \sum_{\chi \in \widehat{G}^H} \widehat{f}(\chi),$$

and so

$$\frac{1}{|H|} \sum_{y \in H} f(y) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}^H} \widehat{f}(\chi).$$

This is the Poisson summation formula.

Exercises

In these exercises, G is a finite abelian group of order n.

1. Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times n$ matrices. Prove that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
2. Define the matrices R and S by (4.12) and (4.13). Prove that $S = R^{-1}$.
3. Let $G = \{x_1, \ldots, x_n\}$. To every matrix $A = (a_{ij}) \in M_n(\mathbf{C})$ we associate a function $K_A \in L^2(G \times G)$ by $K_A(x_i, x_j) = a_{ij}$. Prove that the map $A \mapsto K_A$ is a vector space isomorphism of $M_n(\mathbf{C})$ onto $L^2(G \times G)$.
4. For $a \in G$ and $h \in L^2(G)$, we have operators T_a and C_h on $L^2(G)$, where T_a is translation by a and C_h is convolution by h. Prove that
$$C_h(\delta_a) = T_a(h).$$

4.6 Gauss Sums and Quadratic Reciprocity

Let m be a positive integer, and $\mathbf{Z}/m\mathbf{Z}$ the ring of congruence classes modulo m. An *additive character modulo* m is a character of the additive group $\mathbf{Z}/m\mathbf{Z}$. Since this group is cyclic, the additive characters are the functions ψ_a defined by

$$\psi_a(k + m\mathbf{Z}) = e^{2\pi i a k/m} = e_m(ak)$$

for $a = 0, 1, \ldots, m-1$, and the map from $\mathbf{Z}/m\mathbf{Z}$ to $\widehat{\mathbf{Z}/m\mathbf{Z}}$ that sends the congruence class $a + m\mathbf{Z}$ to the character ψ_a is an isomorphism of additive groups.

A *multiplicative character modulo* m is a character of the multiplicative group of units $(\mathbf{Z}/m\mathbf{Z})^\times$. The *principal character* χ_0 is defined by $\chi_0(a +$

$m\mathbf{Z}) = 1$ if $(a, m) = 1$. If χ is a multiplicative character of $\mathbf{Z}/m\mathbf{Z}$, then we extend χ to a function on $\mathbf{Z}/m\mathbf{Z}$ by defining $\chi(a + m\mathbf{Z}) = 0$ if $(a, m) \neq 1$. Then $\chi \in L^2(\mathbf{Z}/m\mathbf{Z})$. The Fourier transform of χ is $\widehat{\chi} \in L^2\left(\widehat{\mathbf{Z}/m\mathbf{Z}}\right)$, where

$$\begin{aligned}
\widehat{\chi}(\psi_a) &= \sum_{k+m\mathbf{Z}\in\mathbf{Z}/m\mathbf{Z}} \chi(k + m\mathbf{Z})\overline{\psi_a}(k + m\mathbf{Z}) \\
&= \sum_{\substack{k=1 \\ (k,m)=1}}^{m-1} \chi(k + m\mathbf{Z})e_m(-ak).
\end{aligned}$$

For every integer a and multiplicative character χ, we define the *Gauss sum* $\tau(\chi, a)$ as the Fourier transform of χ evaluated at the additive character ψ_{-a}, that is,

$$\begin{aligned}
\tau(\chi, a) &= \widehat{\chi}(\psi_{-a}) = \sum_{\substack{k=1 \\ (k,m)=1}}^{m-1} \chi(k + m\mathbf{Z})e_m(ak) &\quad (4.17)\\
&= \sum_{k=0}^{m-1} \chi(k + m\mathbf{Z})e_m(ak). &\quad (4.18)
\end{aligned}$$

In this section we study multiplicative characters and Gauss sums only for odd prime moduli p.

Theorem 4.16 *Let χ be a nonprincipal multiplicative character modulo the odd prime p. Then*

$$\tau(\chi, a) = \overline{\chi}(a + p\mathbf{Z})\tau(\chi, 1).$$

Proof. If p divides a, then $e_p(ak) = 1$ for all k, and

$$\tau(\chi, a) = \sum_{k=1}^{p-1} \chi(k + p\mathbf{Z})e_p(ak) = \sum_{k=1}^{p-1} \chi(k + p\mathbf{Z}) = 0$$

by the orthogonality relations (Theorem 4.6).

If p does not divide a, then $|\chi(a+p\mathbf{Z})| = 1$, the set $\{ak : k = 1, \ldots, p-1\}$ is a reduced set of residues modulo p, and

$$\begin{aligned}
\tau(\chi, a) &= \sum_{k=1}^{p-1} \chi(k + p\mathbf{Z})e_p(ak) \\
&= \sum_{k=1}^{p-1} \overline{\chi}(a + p\mathbf{Z})\chi(a + p\mathbf{Z})\chi(k + p\mathbf{Z})e_p(ak) \\
&= \overline{\chi}(a + p\mathbf{Z}) \sum_{k=1}^{p-1} \chi(ak + p\mathbf{Z})e_p(ak)
\end{aligned}$$

$$\begin{aligned} &= \overline{\chi}(a+p\mathbf{Z})\sum_{k=1}^{p-1}\chi(k+p\mathbf{Z})e_p(k) \\ &= \overline{\chi}(a+p\mathbf{Z})\tau(\chi,1). \end{aligned}$$

This completes the proof. □

Let p be an odd prime number, and let $\left(\frac{\cdot}{p}\right)$ be the Legendre symbol modulo p. We define the function $\ell_p \in L^2(\mathbf{Z}/p\mathbf{Z})$ by

$$\ell_p(a+p\mathbf{Z}) = \left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p, \\ 0 & \text{if } p \text{ divides } a. \end{cases}$$

Then ℓ_p is a real-valued multiplicative character of $\mathbf{Z}/p\mathbf{Z}$, and

$$\tau(\ell_p, a) = \widehat{\ell_p}(\psi_{-a}) = \sum_{k=1}^{p-1}\left(\frac{k}{p}\right)e_p(ak).$$

The *classical Gauss sum* is

$$\tau(p) = \tau(\ell_p, 1).$$

By Theorem 4.16,

$$\tau(\ell_p, a) = \left(\frac{a}{p}\right)\tau(p). \tag{4.19}$$

For example,

$$\begin{aligned} \tau(3) &= \tau(\ell_3,1) = \left(\frac{1}{3}\right)e_3(1) + \left(\frac{2}{3}\right)e_3(2) \\ &= e_3(1) - e_3(2) = \left(\frac{-1+i\sqrt{3}}{2}\right) - \left(\frac{-1-i\sqrt{3}}{2}\right) \\ &= i\sqrt{3} \end{aligned}$$

and

$$\tau(\ell_3, 2) = \left(\frac{2}{3}\right)\tau(3) = -i\sqrt{3}.$$

Theorem 4.17 *If p is an odd prime and $(a,p)=1$, then*

$$\tau(\ell_p, a) = \sum_{x=0}^{p-1} e_p(ax^2).$$

In particular,

$$\tau(p) = \sum_{x=0}^{p-1} e^{2\pi i x^2/p}.$$

Proof. The set $R = \{k \in \{1, \ldots, p-1\} : \ell_p(k + p\mathbf{Z}) = 1\}$ is a set of representatives of the congruence classes of quadratic residues modulo p, and $N = \{k \in \{1, \ldots, p-1\} : \ell_p(k+p\mathbf{Z}) = -1\}$ is a set of representatives of the congruence classes of quadratic nonresidues modulo p. We have $|R| = |N| = (p-1)/2$. If $x^2 \equiv k \pmod{p}$, then also $(p-x)^2 \equiv k \pmod{p}$. Let $x \not\equiv 0 \pmod{p}$. Since p is odd, $x \not\equiv p - x \pmod{p}$, and

$$\sum_{x=1}^{p-1} e_p(ax^2) = 2 \sum_{k \in R} e_p(ak).$$

It follows that

$$\begin{aligned}
\tau(\ell_p, a) &= \sum_{k=1}^{p-1} \left(\frac{k}{p}\right) e_p(ak) \\
&= \sum_{k \in R} e_p(ak) - \sum_{k \in N} e_p(ak) \\
&= 2\sum_{k \in R} e_p(ak) - \sum_{k \in R \cup N} e_p(ak) \\
&= 1 + 2\sum_{k \in R} e_p(ak) - \sum_{k=0}^{p-1} e_p(ak) \\
&= 1 + \sum_{x=1}^{p-1} e_p(ax^2) \\
&= \sum_{x=0}^{p-1} e_p(ax^2).
\end{aligned}$$

This completes the proof. □

Theorem 4.18 *If p is prime and $(a, p) = 1$, then*

$$\tau(\ell_p, a)^2 = \left(\frac{-1}{p}\right) p = (-1)^{\frac{p-1}{2}} p.$$

Proof. If p does not divide a, then

$$\begin{aligned}
\tau(\ell_p, a)^2 &= \sum_{x=1}^{p-1} \left(\frac{x}{p}\right) e_p(ax) \sum_{y=1}^{p-1} \left(\frac{y}{p}\right) e_p(ay) \\
&= \sum_{x=1}^{p-1}\sum_{y=1}^{p-1} \left(\frac{xy}{p}\right) e_p(a(x+y)).
\end{aligned}$$

Let $(x,p)=1$. Then $\{x,2x,\ldots,(p-1)x\}$ is a reduced set of residues modulo p, $\left(\frac{x^2}{p}\right)=1$, and

$$\begin{aligned}\sum_{y=1}^{p-1}\left(\frac{xy}{p}\right)e_p(-a(x+y)) &= \sum_{y=1}^{p-1}\left(\frac{x(xy)}{p}\right)e_p(-a(x+xy))\\ &= \sum_{y=1}^{p-1}\left(\frac{x^2y}{p}\right)e_p(-ax(1+y))\\ &= \sum_{y=1}^{p-1}\left(\frac{x^2}{p}\right)\left(\frac{y}{p}\right)e_p(-ax(1+y))\\ &= \sum_{y=1}^{p-1}\left(\frac{y}{p}\right)e_p(-ax(1+y)).\end{aligned}$$

Since

$$\sum_{x=1}^{p-1}e_p(-ax(1+y))=\begin{cases} p-1 & \text{if } y\equiv p-1 \pmod p,\\ -1 & \text{if } y\not\equiv p-1 \pmod p,\end{cases}$$

it follows that

$$\begin{aligned}\tau(\ell_p,a)^2 &= \sum_{x=1}^{p-1}\sum_{y=1}^{p-1}\left(\frac{xy}{p}\right)e_p(a(x+y))\\ &= \sum_{y=1}^{p-1}\left(\frac{y}{p}\right)\sum_{x=1}^{p-1}e_p(-ax(1+y))\\ &= \left(\frac{-1}{p}\right)(p-1)-\sum_{y=1}^{p-2}\left(\frac{y}{p}\right)\\ &= \left(\frac{-1}{p}\right)p-\sum_{y=1}^{p-1}\left(\frac{y}{p}\right)\\ &= \left(\frac{-1}{p}\right)p\\ &= (-1)^{\frac{p-1}{2}}p,\end{aligned}$$

by Theorem 3.14. □

Theorem 4.19 *Let p and q be distinct odd prime numbers. If $(a,p)=1$, then*

$$\tau(\ell_p,a)^{q-1}\equiv(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)\pmod q.$$

Proof. By Theorem 4.18 and Theorem 3.12,

$$\begin{aligned} \tau(\ell_p,a)^{q-1} &= \left(\tau(\ell_p,a)^2\right)^{\frac{q-1}{2}} \\ &= \left((-1)^{\frac{p-1}{2}}p\right)^{\frac{q-1}{2}} \\ &= (-1)^{\frac{p-1}{2}\frac{q-1}{2}}p^{\frac{q-1}{2}} \\ &\equiv (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right) \pmod q. \end{aligned}$$

This completes the proof. □

Recall that if G is a finite abelian group, then the map $\Delta : G \to \widehat{\widehat{G}}$ defined by

$$\Delta(a)(\chi) = \chi(a)$$

is an isomorphism.

Theorem 4.20 *If p and q are distinct odd primes, then*

$$\left(\widehat{\widehat{\ell_p}^{\,q}}\right)(\Delta(-q+p\mathbf{Z})) = p\tau(p)^{q-1}\left(\frac{q}{p}\right).$$

Proof. The function on the left side of the equation is a bit complicated. Let $G = \mathbf{Z}/p\mathbf{Z}$. Since $\ell_p \in L^2(G)$, it follows that the Fourier transform $\widehat{\ell_p} \in L^2(\widehat{G})$, and also its qth power $\widehat{\ell_p}^{\,q} \in L^2\left(\widehat{G}\right)$. The Fourier transform of this function is $\widehat{\widehat{\ell_p}^{\,q}} \in L^2\left(\widehat{\widehat{G}}\right)$, and so its domain is $\widehat{\widehat{G}} = \{\Delta(a+p\mathbf{Z}) : a+p\mathbf{Z} \in G\}$. We have

$$\begin{aligned} \left(\widehat{\widehat{\ell_p}^{\,q}}\right)(\Delta(-q+p\mathbf{Z})) &= \sum_{x=0}^{p-1} \widehat{\ell_p}^{\,q}(\psi_x)\overline{\Delta(-q+p\mathbf{Z})}(\psi_x) \\ &= \sum_{x=0}^{p-1} \left(\widehat{\ell_p}(\psi_x)\right)^q \overline{\Delta(-q+p\mathbf{Z})(\psi_x)} \\ &= \sum_{x=0}^{p-1} \tau(\ell_p,-x)^q\overline{\psi_x(-q+p\mathbf{Z})} \\ &= \sum_{x=1}^{p-1} \left(\left(\frac{-x}{p}\right)\tau(p)\right)^q \psi_x(q+p\mathbf{Z}) \\ &= \tau(p)^q\sum_{x=1}^{p-1}\left(\frac{-x}{p}\right)e_p(qx) \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{-q}{p}\right)\tau(p)^q \sum_{x=1}^{p-1}\left(\frac{qx}{p}\right)e_p(qx) \\
&= \left(\frac{-q}{p}\right)\tau(p)^q \sum_{x=1}^{p-1}\left(\frac{x}{p}\right)e_p(x) \\
&= \left(\frac{-q}{p}\right)\tau(p)^{q+1} \\
&= \left(\frac{-q}{p}\right)\left(\frac{-1}{p}\right)p\tau(p)^{q-1} \\
&= p\tau(p)^{q-1}\left(\frac{q}{p}\right),
\end{aligned}$$

by Theorem 4.18. This completes the proof. □

Theorem 4.21 *If p and q are distinct odd primes, then*

$$\left(\widehat{\widehat{\ell_p}^{\,q}}\right)(\Delta(-q+p\mathbf{Z})) = p \sum_{\substack{x_1+\ldots+x_q\equiv q \pmod p \\ 1\le x_i\le p-1}} \left(\frac{x_1\cdots x_q}{p}\right).$$

Proof. Let k be a positive integer. By Exercise 10 in Section 4.3, a product of Fourier transforms is the Fourier transform of the convolution, and so

$$\widehat{\ell_p}^{\,k} = \underbrace{\widehat{\ell_p}*\cdots*\widehat{\ell_p}}_{k\text{ times}} = \underbrace{\widehat{\ell_p*\cdots*\ell_p}}_{k\text{ times}}.$$

By (4.8) of Theorem 4.8, for every integer a we have

$$\begin{aligned}
\left(\widehat{\widehat{\ell_p}^{\,k}}\right)(\Delta(-a+p\mathbf{Z})) &= \underbrace{\widehat{\widehat{\ell_p*\cdots*\ell_p}}}_{k\text{ times}}(\Delta(-a+p\mathbf{Z})) \\
&= p\underbrace{\ell_p*\cdots*\ell_p}_{k\text{ times}}(a+p\mathbf{Z}).
\end{aligned}$$

By Exercise 12 in Section 4.3,

$$\underbrace{\ell_p*\cdots*\ell_p}_{k\text{ times}}(a+p\mathbf{Z}) = \sum_{\substack{x_1+\cdots+x_k\equiv a \pmod p \\ 1\le x_i\le p-1}} \left(\frac{x_1\cdots x_k}{p}\right).$$

If $k=a=q$, then

$$\left(\widehat{\widehat{\ell_p}^{\,q}}\right)(\Delta(-q+p\mathbf{Z})) = p \sum_{\substack{x_1+\cdots+x_q\equiv q \pmod p \\ 1\le x_i\le p-1}} \left(\frac{x_1\cdots x_q}{p}\right).$$

This completes the proof. □

We can now give a second proof of the quadratic reciprocity law. Let p and q be distinct odd primes. By Theorem 4.20 and Theorem 4.21,

$$p\tau(p)^{q-1}\left(\frac{q}{p}\right) = p \sum_{\substack{x_1+\ldots+x_q\equiv q \pmod p \\ 1\le x_i\le p-1}} \left(\frac{x_1\cdots x_q}{p}\right).$$

By Exercise 14 in Section 3.4,

$$\sum_{\substack{x_1+\cdots+x_q\equiv q \pmod p \\ 1\le x_i\le p-1}} \left(\frac{x_1\cdots x_q}{p}\right) \equiv 1 \pmod q,$$

and so

$$\tau(p)^{q-1}\left(\frac{q}{p}\right) \equiv 1 \pmod q.$$

By Theorem 4.19,

$$\tau(p)^{q-1} \equiv (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right) \pmod q,$$

and so

$$(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) \equiv 1 \pmod q.$$

It follows that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

This is the quadratic reciprocity law.

Exercises

1. Show that
$$\tau(5) = 2\left(\cos\frac{\pi}{5} + \cos\frac{2\pi}{5}\right).$$

2. Show that
$$\tau(7) = i2\left(\sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} - \sin\frac{\pi}{7}\right).$$

3. Let p be an odd prime and χ_0 the principal character modulo p. Prove that if p divides a, then $\tau(a,\chi_0) = p-1$.

4. Let g be a primitive root modulo the prime p. Prove that, for every integer b, the function χ_b defined by

$$\chi_b(g^j + p\mathbf{Z}) = e^{2\pi i b j/(p-1)} = e_{p-1}(bj) \tag{4.20}$$

is a multiplicative character modulo p.

Hint: Every congruence class in $(\mathbf{Z}/p\mathbf{Z})^\times$ is uniquely of the form $g^j + p\mathbf{Z}$ for $j = 0, 1, \ldots, p-2$, and the map from $(\mathbf{Z}/p\mathbf{Z})^\times$ to $\mathbf{Z}/(p-1)\mathbf{Z}$ defined by $g^j + p\mathbf{Z} \mapsto j + (p-1)\mathbf{Z}$ is an isomorphism.

5. Prove that the dual group of $(\mathbf{Z}/p\mathbf{Z})^\times$ is the set of functions χ_b defined by (4.20) for $b = 0, 1, \ldots, p-2$.

6. Prove that

$$\chi_b^{-1} = \overline{\chi_b} = \chi_{p-1-b}$$

for $b = 0, 1, \ldots, p-2$.

7. Prove that

$$\chi_b(-1 + p\mathbf{Z}) = (-1)^b$$

for $b = 0, 1, \ldots, p-2$.

8. Let p be an odd prime number, and g a primitive root modulo p. Define the multiplicative characters χ_b by (4.20). Prove that

$$\ell_p = \chi_{(p-1)/2}.$$

9. Let χ be a multiplicative character modulo m, and let a and b be integers relatively prime to m. Prove that

$$\chi(a)\widehat{\chi}(\psi_a) = \chi(b)\widehat{\chi}(\psi_b).$$

10. Let χ be a multiplicative character modulo m. Prove that

$$\chi = \frac{1}{m}\sum_{a=0}^{m-1} \tau(\chi, -a)\psi_a.$$

11. Let ψ be an additive character modulo m and χ a multiplicative character modulo m. Prove that

$$\widehat{\chi}(\psi^{-1}) = \widehat{\psi}(\chi^{-1}).$$

4.7 The Sign of the Gauss Sum

For the odd prime number p, we consider the Gauss sum

$$\tau(p) = \tau(\ell_p, 1) = \sum_{k=1}^{p-1} \left(\frac{k}{p}\right) e_p(k) = \sum_{x=0}^{p-1} e^{2\pi i x^2/p}.$$

By Theorem 4.18,

$$\tau(p)^2 = \begin{cases} p & \text{if } p \equiv 1 \pmod 4, \\ -p & \text{if } p \equiv 3 \pmod 4, \end{cases}$$

and so

$$\tau(p) = \begin{cases} \pm\sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ \pm i\sqrt{p} & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

In this section we determine the sign of $\tau(p)$. We shall prove that

$$\tau(p) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ i\sqrt{p} & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

Recall that for the cyclic group $G = \mathbf{Z}/n\mathbf{Z}$ of order n, the character group $\widehat{G}$ consists of all functions of the form

$$\psi_a(x + n\mathbf{Z}) = e_n(ax).$$

Moreover, the map from G to $\widehat{G}$ defined by $a + n\mathbf{Z} \mapsto \psi_a$ is a group isomorphism. If $\lambda \in L^2(\widehat{G})$, then there is a function $\lambda^\sharp \in L^2(G)$ defined by

$$\lambda^\sharp(a + n\mathbf{Z}) = \lambda(\psi_a).$$

The map $\lambda \mapsto \lambda^\sharp$ is a vector space isomorphism from $L^2(\widehat{G})$ onto $L^2(G)$. The Fourier transform is a a vector space isomorphism from $L^2(G)$ onto $L^2(\widehat{G})$. Define $\mathcal{F} : L^2(G) \to L^2(G)$ as the composition of the Fourier transform with the $\sharp$ map. If $f \in L^2(G)$, then

$$\begin{aligned} \mathcal{F}(f)(a + n\mathbf{Z}) &= \left(\widehat{f}\right)^\sharp (a + n\mathbf{Z}) \\ &= \widehat{f}(\psi_a) \\ &= \sum_{x=0}^{n-1} f(x + n\mathbf{Z})\overline{\psi_a}(x + n\mathbf{Z}) \\ &= \sum_{x=0}^{n-1} f(x + n\mathbf{Z})\omega^{-ax}, \end{aligned}$$

where

$$\omega = e_n(1) = e^{2\pi i/n}.$$

The linear operator $\mathcal{F}$ is also called the Fourier transform.

Theorem 4.22 *For all functions $f \in L^2(\mathbf{Z}/n\mathbf{Z})$,*

$$\mathcal{F}^2(f)(a+n\mathbf{Z}) = nf(-a+n\mathbf{Z}).$$

Proof. This is similar to the proof of (4.8) in Theorem 4.8. Writing $\mathcal{F}(f) = g$, we have

$$g(x+n\mathbf{Z}) = \sum_{y=0}^{n-1} f(y+n\mathbf{Z})\omega^{-xy}$$

and

$$\begin{aligned}
\mathcal{F}^2(f)(a+n\mathbf{Z}) &= \mathcal{F}(g)(a+n\mathbf{Z}) \\
&= \sum_{x=0}^{n-1} g(x+n\mathbf{Z})\omega^{-ax} \\
&= \sum_{x=0}^{n-1}\sum_{y=0}^{n-1} f(y+n\mathbf{Z})\omega^{-xy}\omega^{-ax} \\
&= \sum_{y=0}^{n-1} f(y+n\mathbf{Z}) \sum_{x=0}^{n-1} \omega^{-x(a+y)} \\
&= nf(-a+n\mathbf{Z}).
\end{aligned}$$

This completes the proof. □

The vector space $L^2(G)$ has a basis $\{\delta_k\}_{k=0}^{n-1}$, where the delta function δ_k is defined by

$$\delta_k(x+n\mathbf{Z}) = \begin{cases} 1 & \text{if } x \equiv k \pmod{n}, \\ 0 & \text{if } x \not\equiv k \pmod{n}. \end{cases}$$

We shall compute the matrix of the linear operator $\mathcal{F}$ with respect to this basis. We have

$$\mathcal{F}(\delta_k)(j+n\mathbf{Z}) = \sum_{x=0}^{n-1} \delta_k(x+n\mathbf{Z})\omega^{-jx} = \omega^{-jk},$$

and so

$$\mathcal{F}(\delta_k) = \sum_{j=0}^{n-1} \omega^{-jk}\delta_j.$$

Therefore, the matrix of $\mathcal{F}$ with respect to the basis $\{\delta_k\}_{k=0}^{n-1}$ is

$$M(\mathcal{F}) = \left(\omega^{-jk}\right)_{j,k=0}^{n-1}. \tag{4.21}$$

For any positive integer n we define the Gauss sum

$$\tau(n) = \sum_{k=0}^{n-1} e^{2\pi i k^2/n}.$$

By Theorem 4.17, this is consistent with our previous definition of $\tau(p)$ for p prime. Since $\overline{\omega^{-k}} = \omega^k$ for all integers k, it follows that the trace of the matrix $M(\mathcal{F})$ is

$$\mathrm{tr}(M(\mathcal{F})) = \sum_{k=0}^{n-1} \omega(-k^2) = \overline{\sum_{k=0}^{n-1} \omega(k^2)} = \overline{\tau(n)}.$$

Since the determinant and trace of a linear operator on a finite-dimensional vector space are independent of the choice of basis for the vector space, it follows that the trace of the Fourier transform $\mathcal{F}$ on the group $\mathbf{Z}/n\mathbf{Z}$ is the complex conjugate of the Gauss sum $\tau(n)$.

Theorem 4.23 *Let n be an odd positive integer and $G = \mathbf{Z}/n\mathbf{Z}$ the cyclic group of order n. Then the determinant of the Fourier transform $\mathcal{F}$ on $L^2(G)$ is*

$$\det(\mathcal{F}) = \begin{cases} (-1)^k n^{n/2} & \text{if } n = 4k+1, \\ (-1)^k i n^{n/2} & \text{if } n = 4k+3. \end{cases}$$

Proof. We shall compute the determinant of the matrix $M(\mathcal{F})$ in two ways. Let $\omega = e^{2\pi i/n}$. The square of $M(\mathcal{F})$ is the matrix $B = (b_{jk})_{j,k=0}^{n-1}$, where

$$b_{jk} = \sum_{\ell=0}^{n-1} \omega^{-j\ell}\omega^{-\ell k} = \sum_{\ell=0}^{n-1} \omega^{-(j+k)\ell} = \begin{cases} n & \text{if } j+k \equiv 0 \pmod{n}, \\ 0 & \text{if } j+k \not\equiv 0 \pmod{n}, \end{cases}$$

and so (by Exercise 4)

$$\det(M(\mathcal{F}))^2 = \det(B) = (-1)^{(n-1)/2} n^n = i^{n-1} n^n.$$

Then

$$\det(M(\mathcal{F})) = \pm i^{(n-1)/2} n^{n/2}. \tag{4.22}$$

The determinant of $M(\mathcal{F})$ is also a *Vandermonde determinant* (Nathanson [103, pp. 78–81]), whose value is

$$\begin{aligned} \det(\mathcal{F}) &= \prod_{0 \le j < k \le n-1} \left(\omega^{-k} - \omega^{-j}\right) \\ &= \prod_{0 \le j < k \le n-1} \omega^{-(j+k)/2} \left(\omega^{-(k-j)/2} - \omega^{(k-j)/2}\right) \\ &= \prod_{0 \le j < k \le n-1} \omega^{-(j+k)/2} \left(-2i \sin\left(\frac{(k-j)\pi}{n}\right)\right) \end{aligned}$$

$$
\begin{aligned}
&= \prod_{0\le j<k\le n-1} \omega^{-(j+k)/2} \prod_{0\le j<k\le n-1} \left(-2i\sin\left(\frac{(k-j)\pi}{n}\right)\right) \\
&= \omega^{-\sum_{0\le j<k\le n-1}(j+k)/2}(-i)^{n(n-1)/2} \prod_{0\le j<k\le n-1} 2\sin\left(\frac{(k-j)\pi}{n}\right).
\end{aligned}
$$

We can compute the exponent of ω as follows:

$$
\begin{aligned}
\sum_{0\le j<k\le n-1} \frac{j+k}{2} &= \frac{1}{2}\sum_{k=1}^{n-1}\sum_{j=0}^{k-1}(j+k) \\
&= \frac{1}{2}\sum_{k=1}^{n-1}\left(\frac{k(k-1)}{2}+k^2\right) \\
&= \frac{1}{4}\sum_{k=1}^{n-1}\left(3k^2-k\right) \\
&= n\left(\frac{n-1}{2}\right)^2,
\end{aligned}
$$

by Exercise 6. Since n is odd, it follows that

$$
\sum_{0\le j<k\le n-1} \frac{j+k}{2} \equiv 0 \pmod{n},
$$

and so

$$
\omega^{-\sum_{0\le j<k\le n-1}(j+k)/2} = 1.
$$

If $0 \le j < k \le n-1$, then $0 < \frac{(k-j)\pi}{n}) < \pi$ and $\sin\left(\frac{(k-j)\pi}{n}\right) > 0$. Therefore,

$$
\det(M(\mathcal{F})) = (-i)^{n(n-1)/2} \prod_{0\le j<k\le n-1} 2\sin\left(\frac{(k-j)\pi}{n}\right), \tag{4.23}
$$

where

$$
\prod_{0\le j<k\le n-1} 2\sin\left(\frac{(k-j)\pi}{n}\right) > 0.
$$

Comparing (4.22) and (4.23), we obtain

$$
\det(\mathcal{F}) = (-i)^{n(n-1)/2} n^{n/2}.
$$

By Exercise 7,

$$
(-i)^{n(n-1)/2} = \begin{cases} (-1)^k & \text{if } n = 4k+1, \\ (-1)^k i & \text{if } n = 4k+3. \end{cases}
$$

This completes the proof. □

Theorem 4.24 *Let p be an odd prime and $G = \mathbf{Z}/p\mathbf{Z}$ the cyclic group of order p. Then the determinant of the Fourier transform $\mathcal{F}$ on $L^2(G)$ is*

$$\det(\mathcal{F}) = p \prod_{b=1}^{p-2} \tau(\chi_b, 1),$$

where χ_b is the multiplicative character modulo p defined by (4.20) for $b = 0, 1, \ldots, p-2$.

Proof. The $p-1$ functions $\chi_0, \chi_1, \ldots, \chi_{p-2}$ are orthogonal in $L^2(G)$, since

$$(\chi_a, \chi_b) = \sum_{x=0}^{p-1} \chi_a(x + p\mathbf{Z})\overline{\chi_b}(x + p\mathbf{Z}) = \begin{cases} p-1 & \text{if } a = b, \\ 0 & \text{if } a \neq b \end{cases}$$

by Theorem 4.7. Let δ_0 be the delta function at 0, that is,

$$\delta_0(x + p\mathbf{Z}) = \begin{cases} 1 & \text{if } x \equiv 0 \pmod{p}, \\ 0 & \text{if } x \not\equiv 0 \pmod{p}. \end{cases}$$

Then

$$(\delta_0, \delta_0) = 1$$

and

$$(\chi_b, \delta_0) = \sum_{x=0}^{p-1} \chi_b(x + p\mathbf{Z})\delta_0(x + p\mathbf{Z}) = \chi_b(p\mathbf{Z}) = 0.$$

It follows that the set $\{\delta_0, \chi_0, \chi_1, \ldots, \chi_{p-2}\}$ is an orthogonal set of p functions in $L^2(G)$, and so is a basis for $L^2(G)$. This basis is called the *basis of multiplicative characters* for $L^2(G)$. We shall compute the matrix of the Fourier transform $\mathcal{F}$ with respect to this basis.

For every congruence class $a + p\mathbf{Z} \in G$ we have

$$\begin{aligned} \mathcal{F}(\delta_0)(a + p\mathbf{Z}) &= \widehat{\delta_0}(\psi_a) \\ &= \sum_{x=0}^{p-1} \delta_0(x + p\mathbf{Z})\overline{\psi_a}(x + p\mathbf{Z}) \\ &= \overline{\psi_a}(p\mathbf{Z}) \\ &= 1 \\ &= \delta_0(a + p\mathbf{Z}) + \chi_0(a + p\mathbf{Z}), \end{aligned}$$

where χ_0 is the principal multiplicative character modulo p. Therefore,

$$\mathcal{F}(\delta_0) = \delta_0 + \chi_0.$$

Similarly,

$$\begin{aligned}
\mathcal{F}(\chi_0)(a+p\mathbf{Z}) &= \widehat{\chi_0}(\psi_a) \\
&= \sum_{x=0}^{p-1} \chi_0(x+p\mathbf{Z})\overline{\psi_a}(x+p\mathbf{Z}) \\
&= \sum_{x=1}^{p-1} \overline{\psi_a}(x+p\mathbf{Z}) \\
&= \sum_{x=0}^{p-1} \psi_{-a}(x+p\mathbf{Z}) - 1 \\
&= \begin{cases} p-1 & \text{if } a \equiv 0 \pmod{p} \\ -1 & \text{if } a \not\equiv 0 \pmod{p} \end{cases} \\
&= (p-1)\delta_0(a+p\mathbf{Z}) - \chi_0(a+p\mathbf{Z}),
\end{aligned}$$

and so

$$\mathcal{F}(\chi_0) = (p-1)\delta_0 - \chi_0.$$

By Theorem 4.16, and by Exercises 6 and 7 in Section 4.6, if $b \not\equiv 0 \pmod{p-1}$, then

$$\begin{aligned}
\mathcal{F}(\chi_b)(a+p\mathbf{Z}) &= \widehat{\chi_b}(\psi_a) \\
&= \tau(\chi_b, -a) \\
&= \tau(\chi_b, 1)\overline{\chi_b}(-a+p\mathbf{Z}) \\
&= \tau(\chi_b, 1)\chi_{p-1-b}(-a+p\mathbf{Z}) \\
&= (-1)^b \tau(\chi_b, 1)\chi_{p-1-b}(a+p\mathbf{Z}),
\end{aligned}$$

and so

$$\mathcal{F}(\chi_b) = (-1)^b \tau(\chi_b, 1)\chi_{p-1-b}. \tag{4.24}$$

This determines the matrix of $\mathcal{F}$ with respect to the basis of multiplicative characters. For example, if $p = 5$, this matrix is

$$\begin{pmatrix}
1 & 4 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\tau(\chi_3, 1) \\
0 & 0 & 0 & \tau(\chi_2, 1) & 0 \\
0 & 0 & -\tau(\chi_1, 1) & 0 & 0
\end{pmatrix}.$$

By Exercise 4, the determinant of this matrix is

$$\begin{aligned}
\det(\mathcal{F}) &= -p(-1)^{(p-3)/2} \prod_{b=1}^{p-2} (-1)^b \tau(\chi_b, 1) \\
&= p(-1)^{(p-1)/2} \prod_{b=1}^{p-2} (-1)^b \prod_{b=1}^{p-2} \tau(\chi_b, 1)
\end{aligned}$$

$$= p\prod_{b=1}^{p-2}\tau(\chi_b,1).$$

This completes the proof. □

We can now determine the sign of the classical Gaussian sum.

Theorem 4.25 *If p is an odd prime, then*

$$\tau(p)=\sum_{x=1}^{p-1}e^{2\pi i x^2/p}=\begin{cases}\sqrt{p} & \text{if } p\equiv 1 \pmod 4,\\ i\sqrt{p} & \text{if } p\equiv 3 \pmod 4.\end{cases}$$

Proof. By (4.24), we have

$$\mathcal{F}(\chi_b)=(-1)^b\tau(\chi_b,1)\chi_{p-1-b}$$

and so

$$\begin{aligned}\mathcal{F}^2(\chi_b) &= \mathcal{F}\left((-1)^b\tau(\chi_b,1)\chi_{p-1-b}\right)\\ &= (-1)^b\tau(\chi_b,1)\mathcal{F}\left(\chi_{p-1-b}\right)\\ &= (-1)^b\tau(\chi_b,1)(-1)^{p-1-b}\tau(\chi_{p-1-b},1)\chi_b\\ &= \tau(\chi_b,1)\tau(\chi_{p-1-b},1)\chi_b.\end{aligned}$$

On the other hand, applying Fourier inversion (Theorem 4.22), we obtain

$$\begin{aligned}\mathcal{F}^2(\chi_b)(a+p\mathbf{Z}) &= p\chi_b(-a+p\mathbf{Z})\\ &= \chi_b(-1+p\mathbf{Z})p\chi_b(a+p\mathbf{Z})\\ &= (-1)^b p\chi_b(a+p\mathbf{Z}),\end{aligned}$$

and so

$$\mathcal{F}^2(\chi_b)=(-1)^b p\chi_b.$$

It follows that

$$\tau(\chi_b,1)\tau(\chi_{p-1-b},1)=(-1)^b p.$$

Let $r=(p-1)/2$. It follows from Exercise 8 in Section 4.6 that $\ell_p=\chi_r$ and $\tau(p)=\tau(\chi_r,1)$. By Theorem 4.24,

$$\begin{aligned}\det(\mathcal{F}) &= p\prod_{b=1}^{p-2}\tau(\chi_b,1)\\ &= p\tau(p)\prod_{b=1}^{r-1}\tau(\chi_b,1)\tau(\chi_{p-1-b},1)\\ &= p\tau(p)\prod_{b=1}^{r-1}\left((-1)^b p\right)\\ &= (-1)^{r(r-1)/2}p^{(p-1)/2}\tau(p).\end{aligned}$$

By Theorem 4.23,

$$\det(\mathcal{F}) = \begin{cases} (-1)^k p^{p/2} & \text{if } p = 4k+1, \\ (-1)^k i p^{p/2} & \text{if } p = 4k+3. \end{cases}$$

If $p = 4k+1$, then $r = 2k$ and

$$\begin{aligned} (-1)^{r(r-1)/2} p^{(p-1)/2} \tau(p) &= (-1)^{k(2k-1)} p^{(p-1)/2} \tau(p) \\ &= (-1)^k p^{(p-1)/2} \tau(p) \\ &= (-1)^k p^{p/2}, \end{aligned}$$

and so

$$\tau(p) = \sqrt{p}.$$

If $p = 4k+3$, then $r = 2k+1$ and

$$\begin{aligned} (-1)^{r(r-1)/2} p^{(p-1)/2} \tau(p) &= (-1)^{k(2k+1)} p^{(p-1)/2} \tau(p) \\ &= (-1)^k p^{(p-1)/2} \tau(p) \\ &= (-1)^k i p^{p/2}, \end{aligned}$$

and so

$$\tau(p) = i\sqrt{p}.$$

This completes the proof. □

Exercises

1. Prove that

$$2\left(\cos\frac{\pi}{5} + \cos\frac{2\pi}{5}\right) = \sqrt{5}.$$

and

$$2\left(\sin\frac{2\pi}{7} + \sin\frac{4\pi}{7} - \sin\frac{\pi}{7}\right) = \sqrt{7}.$$

Hint: Consider the Gauss sums $\tau(5)$ and $\tau(7)$.

2. Prove that

$$\tau(\ell_p, a) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4 \text{ and } \left(\frac{a}{p}\right) = 1, \\ -\sqrt{p} & \text{if } p \equiv 1 \pmod 4 \text{ and } \left(\frac{a}{p}\right) = -1, \\ i\sqrt{p} & \text{if } p \equiv 3 \pmod 4 \text{ and } \left(\frac{a}{p}\right) = 1, \\ -i\sqrt{p} & \text{if } p \equiv 3 \pmod 4 \text{ and } \left(\frac{a}{p}\right) = -1. \end{cases}$$

3. Let $\omega = e^{2\pi i/3}$. Compute the trace and determinant of the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{pmatrix}$$

4. Let $A = (a_{j,k})_{j,k=1}^{n-1}$ be an $n-1 \times n-1$ matrix such that $a_{j,k} = 0$ if $j + k \not\equiv 0 \pmod{n}$. For example, if $n = 4$, then

$$\begin{pmatrix} 0 & 0 & a_{1,3} \\ 0 & a_{2,2} & 0 \\ a_{3,1} & 0 & 0 \end{pmatrix}.$$

Prove that

$$\det(A) = \begin{cases} (-1)^{(n-1)/2} \prod_{j=1}^{n-1} a_{j,n-j} & \text{if } n \text{ is odd,} \\ (-1)^{(n-2)/2} \prod_{j=1}^{n-1} a_{j,n-j} & \text{if } n \text{ is even.} \end{cases}$$

Let $B = (b_{j,k})_{j,k=0}^{n-1}$ be an $n \times n$ matrix such that $b_{j,k} = n$ if $j + k \equiv 0 \pmod{n}$ and $b_{j,k} = 0$ if $j + k \not\equiv 0 \pmod{n}$. For example, if $n = 4$, then

$$\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}.$$

Prove that

$$\det(B) = \begin{cases} (-1)^{(n-1)/2} n^n & \text{if } n \text{ is odd,} \\ (-1)^{(n-2)/2} n^n & \text{if } n \text{ is even.} \end{cases}$$

5. Let I_n denote the $n \times n$ identity matrix. Prove that $M(\mathcal{F})^4 = n^2 I_n$ and so

$$\det(\mathcal{F})^4 = n^{2n}.$$

6. Prove that for every positive integer n,

$$\sum_{k=1}^{n-1} \left(3k^2 - k\right) = n\,(n-1)^2\,.$$

7. Let n be an odd integer. Prove that

$$(-i)^{n(n-1)/2} = \begin{cases} (-1)^k & \text{if } n = 4k+1, \\ (-1)^k i & \text{if } n = 4k+3. \end{cases}$$

8. Prove that the Legendre symbol is an eigenvector of the Fourier transform with eigenvalue $(-1)^{(p-1)/2}\tau(p)$.

Hint: Exercise 8 in Section 4.6.

4.8 Notes

A comprehensive survey of analysis and trace formulae on finite abelian and nonabelian groups is Terras, *Fourier Analysis on Finite Groups and Applications* [141]. Our proof of the sign of the Gauss sum uses an argument of Schur [126] that appears Landau [87, pp. 207–212] and Auslander and Tolimieri [7]. See Berndt and Evans [8] for a review of Gauss sums, and Berndt, Evans, and Williams, *Gauss and Jacobi Sums* [9] for an exhaustive monograph.

For much more sophisticated studies of harmonic analysis in algebraic number theory, see Ramakrishnan and Valenza, *Fourier Analysis on Number Fields* [120], and Weil's classic *Basic Number Theory* [154].

5

The *abc* Conjecture

5.1 Ideals and Radicals

In this chapter a ring is always a commutative ring with identity. An additive subgroup I of a ring R is called an *ideal* if $ar \in I$ for every $a \in I$ and $r \in R$. Both R and $\{0\}$ are ideals in R. The set of even integers is an ideal in the ring $\mathbf{Z}$. Indeed, every additive subgroup of $\mathbf{Z}$ is an ideal in $\mathbf{Z}$. The set of polynomials with constant term equal to 0 is an ideal in the ring $R[t]$ of polynomials with coefficients in the ring R. The intersection of a family of ideals is an ideal (Exercise 19 in Section 3.1).

If A is a nonempty subset of the ring R, then the set of all finite linear combinations of the form $a_1r_1 + \cdots + a_kr_k$ with $a_i \in A$ and $r_i \in R$ is an ideal of R, denoted by $\langle A \rangle$ and called the ideal generated by the set A. An ideal generated by one element $a \in R$ is called a *principal ideal* and denoted by

$$\langle a \rangle = aR = \{ar : r \in R\}.$$

A *principal ring* is a ring in which every ideal is principal. For example, $\mathbf{Z}$ is a principal ring by Theorem 1.3, and $\mathbf{Z}/m\mathbf{Z}$ is a principal ring by Theorem 5.2.

An ideal I in the ring R is called a *prime ideal* if $I \neq R$ and $ab \in I$ implies $a \in I$ or $b \in I$ for all $a, b \in R$. The *spectrum* of the ring R, denoted by $\mathrm{Spec}(R)$, is the set of all prime ideals of R.

Theorem 5.1 *The spectrum of the ring of integers is*

$$\mathit{Spec}(\mathbf{Z}) = \{p\mathbf{Z} : p \textit{ is prime or } p = 0\}.$$

Proof. Since $\mathbf{Z}$ is principal, every ideal is of the form $d\mathbf{Z}$ for some nonnegative integer d. If $d = 0$, then $d\mathbf{Z} = \{0\}$, and the zero ideal is prime, since $ab = 0$ if and only if $a = 0$ or $b = 0$. Let $d \geq 1$. If $d = p$ is prime and $ab \in p\mathbf{Z}$, then p divides ab. By Euclid's lemma, p divides a or p divides b, and so $a \in p\mathbf{Z}$ or $b \in p\mathbf{Z}$. Therefore, $p\mathbf{Z}$ is a prime ideal for every prime number p.

If d is composite, then we can write $d = ab$, where $1 < a \leq b < d$. If $a \in d\mathbf{Z}$, then $a = dk = abk$ for some positive integer k, and so $1 = bk$, which is absurd. Therefore, $a \notin d\mathbf{Z}$ and, similarly, $b \notin d\mathbf{Z}$. Since $d = ab \in d\mathbf{Z}$, it follows that $d\mathbf{Z}$ is not a prime ideal. Thus, the prime ideals in the ring $\mathbf{Z}$ are the ideals of the form $p\mathbf{Z}$, where p is a prime number or $p = 0$. □

An element x in a ring R is called *nilpotent* if there exists a positive integer k such that $x^k = 0$. For example, the additive identity 0 is a nilpotent element of every ring, and the multiplicative identity 1 is never nilpotent. The congruence class $6+27\mathbf{Z}$ is a nilpotent element in the ring $\mathbf{Z}/27\mathbf{Z}$. The set of all nilpotent elements in R is called the *radical* of the ring R, and denoted by $\mathcal{N}(R)$. Thus, the radical of the ring $\mathbf{Z}$ is $\{0\}$. By Exercise 6, the radical of a ring is a proper ideal in the ring. By Exercise 9, the radical of a ring is the intersection of the prime ideals in the ring.

We shall compute the radical of the ring of congruence classes $\mathbf{Z}/m\mathbf{Z}$. Recall that the *radical* of the nonzero integer m is the product of the distinct prime numbers that divide m, that is,

$$\operatorname{rad}(m) = \prod_{p \mid m} p.$$

For example, $\operatorname{rad}(72) = 6, \operatorname{rad}(30) = 30$, and $\operatorname{rad}(-1) = 1$.

Theorem 5.2 *For $m \geq 2$, let $\mathbf{Z}/m\mathbf{Z}$ be the ring of congruence classes modulo m. Then*

(i) $\mathbf{Z}/m\mathbf{Z}$ is principal, and the ideals of $\mathbf{Z}/m\mathbf{Z}$ are the ideals generated by the congruence classes $d + m\mathbf{Z}$, where d is a divisor of m;

(ii) the prime ideals of $\mathbf{Z}/m\mathbf{Z}$ are the ideals generated by the congruence classes $p + m\mathbf{Z}$, where p is a prime divisor of m; and

(iii) the radical of $\mathbf{Z}/m\mathbf{Z}$ is the ideal generated by the congruence class $\operatorname{rad}(m) + m\mathbf{Z}$.

Proof. Let J be an ideal in the ring $R = \mathbf{Z}/m\mathbf{Z}$. Consider the union of congruence classes

$$I = \bigcup_{a+m\mathbf{Z} \in J} (a + m\mathbf{Z}).$$

The set I is an ideal in $\mathbf{Z}$. Since $\mathbf{Z}$ is principal, $I = d\mathbf{Z}$ for some positive integer $d \in I$. Since $m \in m\mathbf{Z} \subseteq I$, it follows that d is a divisor of m.

Moreover, $d + m\mathbf{Z} \in J$, and so the principal ideal generated by $d + m\mathbf{Z}$ in $\mathbf{Z}/m\mathbf{Z}$ is contained in J. If $a+m\mathbf{Z} \in J$, then $a \in a+m\mathbf{Z} \subseteq I$, and so $a = dr$ for some integer r. It follows that $a + m\mathbf{Z} = (d + m\mathbf{Z})(r + m\mathbf{Z})$ belongs to the principal ideal generated by $d + m\mathbf{Z}$. Therefore, J is the principal ideal generated by $d+m\mathbf{Z}$, and $a+m\mathbf{Z} \in J$ if and only if d divides a. (See Exercise 3 for a different proof.)

Next we compute the spectrum of the ring $\mathbf{Z}/m\mathbf{Z}$. Let J be the principal ideal generated by $d+m\mathbf{Z}$, where d divides m and $d \geq 2$. If $d = p$ is prime and

$$(a + m\mathbf{Z})(b + m\mathbf{Z}) = ab + m\mathbf{Z} \in J,$$

then p divides ab and so p divides a or p divides b, that is, $a + m\mathbf{Z} \in J$ or $b + m\mathbf{Z} \in J$, and J is a prime ideal.

If $d = ab$ is composite, where $1 < a \leq b < d$, then $a+m\mathbf{Z} \notin J$, $b+m\mathbf{Z} \notin J$, but $(a+m\mathbf{Z})(b+m\mathbf{Z}) = d+m\mathbf{Z} \in J$, and so J is not a prime ideal. Thus, the prime ideals of the ring $\mathbf{Z}/m\mathbf{Z}$ are the ideals of the form $p+m\mathbf{Z}$, where p is a prime divisor of m.

Finally, the congruence class $a + m\mathbf{Z}$ is nilpotent in R if and only if $(a + m\mathbf{Z})^k = a^k + m\mathbf{Z} = m\mathbf{Z}$ for some positive integer k. Equivalently, $a+m\mathbf{Z}$ is nilpotent if and only if m divides a^k for some positive integer k. By Theorem 1.13, this is possible if and only if a is divisible by $\mathrm{rad}(m)$, and so $\mathcal{N}(\mathbf{Z}/m\mathbf{Z})$ is the ideal generated by the congruence class $\mathrm{rad}(m) + m\mathbf{Z}$. □

Theorem 5.3 *The ring* $\mathbf{C}[t]$ *of polynomials with coefficients in the field* $\mathbf{C}$ *of complex numbers is a principal ring.*

Proof. This is a special case of Exercise 18 in Section 3.1. □

Let $f(t) \in \mathbf{C}[t]$ be a polynomial of degree n. If $\alpha_1, \ldots, \alpha_r$ are the distinct zeros of $f(t)$, then we can factor $f(t)$ into a product of linear terms of the form $f(t) = c_n \prod_{i=1}^{r}(t - \alpha_i)^{m_i}$, where the leading coefficient $c_n \neq 0$ and $m_1 + \cdots + m_r = n$. The *radical of the polynomial* $f(x)$ is defined by

$$\mathrm{rad}(f) = \prod_{i=1}^{r}(t - \alpha_i).$$

The *zero set* of the polynomial $f(t)$ is the finite set

$$Z(f) = \{\alpha \in \mathbf{C} : f(\alpha) = 0\} = \{\alpha_1, \ldots, \alpha_r\}.$$

Let $N_0(f)$ denote the number of distinct zeros of f, that is, $N_0(f) = |Z(f)| = r$. The degree of the radical of $f(t)$ is the number of distinct zeros of $f(t)$, that is,

$$\deg \mathrm{rad}(f) = N_0(f).$$

Theorem 5.4 *Let $f(t) \in \mathbf{C}[t]$ and $R = \mathbf{C}[t]/I$, where $I = \langle f(t) \rangle$ is the principal ideal generated by $f(t)$. The radical of R is the principal ideal generated by rad$(f) + I$.*

Proof. This follows immediately from the observation that if $f(t)$ and $g(t)$ are polynomials with complex coefficients, then there exists a positive integer k such that $f(t)$ divides $g(t)^k$ if and only if $\mathrm{rad}(f)$ divides $g(t)$. □

Exercises

1. Determine $\mathrm{rad}(3^n)$ and $\mathrm{rad}(n!)$ for all $n \geq 0$.

2. Let m and n be nonzero integers. Prove that $\mathrm{rad}(mn) \leq \mathrm{rad}(m)\mathrm{rad}(n)$. Prove that $\mathrm{rad}(mn) = \mathrm{rad}(m)\mathrm{rad}(n)$ if and only if $(m, n) = 1$.

3. Let $f : R \to S$ be a surjective ring homomorphism. Prove that if the ring R is principal, then the ring S is also principal. Apply this to the map $f : \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$ defined by $f(a) = a + m\mathbf{Z}$.

4. Prove that a unit in a ring $R \neq \{0\}$ is never nilpotent.

5. Let R be an *integral domain*, that is, a ring with the property that if $x_1, x_2 \in R$ and $x_1x_2 = 0$, then $x_1 = 0$ or $x_2 = 0$. Prove that if $x_1, \ldots, x_k \in R$ and $x_1 \cdots x_k = 0$, then $x_i = 0$ for some i. Prove that 0 is the only nilpotent element in an integral domain.

6. Let R be a ring and let $\mathcal{N}(R)$ denote the set of all nilpotent elements in R. Prove that $\mathcal{N}(R)$ is an ideal.

   Hint: Prove that if x is nilpotent, then xr is nilpotent for every $r \in R$. Use the binomial theorem to show that if $x^k = y^\ell = 0$, then $(x + y)^{k+\ell-1} = 0$.

7. Prove that if x is nilpotent, then x is contained in every prime ideal of R, and so
$$\mathcal{N}(R) \subseteq \bigcap_{I \in \mathrm{Spec}(R)} I.$$

8. Prove that if x is not nilpotent, then there exists a prime ideal of R that does not contain x.

   Hint: Let $S = \{x^k : k = 1, 2, \ldots\}$. Let $\mathcal{I}$ be the set of all ideals in R that do not contain any element of S. If x is not nilpotent, then $0 \notin S$ and $\{0\} \in \mathcal{I}$. Use Zorn's lemma to prove that the set $\mathcal{I}$ contains a maximal element I, and that I is a prime ideal in R. such that $I \cap S = \emptyset$.

9. Prove that the radical of the ring R is the intersection of all prime ideals of R, that is,
$$\mathcal{N}(R) = \bigcap_{I \in \mathrm{Spec}(R)} I.$$

10. Let $a_1, \ldots, a_k$ be divisors of m, and let $[a_1, \ldots, a_k]$ be their least common multiple. Let $\langle a_i + m\mathbf{Z}\rangle$ denote the principal ideal generated by the congruence class $a_i + m\mathbf{Z}$ in the ring $R = \mathbf{Z}/m\mathbf{Z}$. Prove that
$$\bigcap_{i=1}^{k} \langle a_i + m\mathbf{Z}\rangle = \langle [a_1, \ldots, a_k] + m\mathbf{Z}\rangle.$$
Hint: Observe that $\langle a_i + m\mathbf{Z}\rangle = a_i\mathbf{Z}$ and apply Exercise 30 in Section 1.4.

11. Use Exercises 9 and 10 to prove that
$$\mathcal{N}(\mathbf{Z}/m\mathbf{Z}) = \langle \mathrm{rad}(m) + m\mathbf{Z}\rangle.$$

12. Let I and J be ideals in a ring R. The *product* IJ is the ideal of R generated by the set of all elements of the form xy with $x \in I$ and $y \in J$. In the ring $\mathbf{Z}$, prove that the product of the principal ideals $a\mathbf{Z}$ and $b\mathbf{Z}$ is the ideal $ab\mathbf{Z}$.

13. Let I and J be ideals in the ring R. We say that I *divides* J if I contains J, that is, $J \subset I$. Prove that if P is a prime ideal in R and if P divides the product ideal IJ, then P divides I or P divides J.

14. Let I and J be ideals in $\mathbf{Z}$. Prove that if I divides J, then there exists an ideal K in $\mathbf{Z}$ such that $IK = J$. Prove that every ideal in $\mathbf{Z}$ is uniquely a product of prime ideals.

5.2 Derivations

A *derivation* on a ring R is a map $D : R \to R$ such that
$$D(x + y) = D(x) + D(y) \tag{5.1}$$
and
$$D(xy) = D(x)y + xD(y) \tag{5.2}$$
for all $x, y \in R$. Condition (5.1) says that D is a homomorphism of the additive group structure of R. Condition (5.2) implies (Exercise 1) that $D(1) = 0$ and that, if $x \in R$ is invertible, then
$$D(x^{-1}) = -\frac{D(x)}{x^2}.$$

Moreover, it follows by induction (Exercise 2) that

$$D(x_1 \cdots x_n) = \sum_{i=1}^{n} x_1 \cdots x_{i-1} D(x_i) x_{i+1} \cdots x_n$$

for all $x_1, \ldots, x_n \in R$.

The next result shows that the derivative is a derivation on a polynomial ring.

Theorem 5.5 *Let R be a ring and $R[t]$ the ring of polynomials with coefficients in R. Define $D : R[t] \to R[t]$ by*

$$D\left(\sum_{i=0}^{m} a_i t^i\right) = \sum_{i=1}^{m} i a_i t^{i-1}.$$

Then D is a derivation on $R[t]$.

Proof. Let $f = f(t) = \sum_{i=0}^{m} a_i t^i$ and $g = g(t) = \sum_{j=0}^{n} b_j t^j$. It is immediate that $D(f+g) = D(f) + D(g)$, and so D is a homomorphism of the additive group of polynomials. Since

$$f(t)g(t) = \sum_{i=0}^{m}\sum_{j=0}^{n} a_i t^i b_j t^j = \sum_{k=0}^{m+n} \sum_{i+j=k} a_i b_j t^k,$$

we have

$$\begin{aligned} D(fg) &= \sum_{k=1}^{m+n} k \sum_{i+j=k} a_i b_j t^{k-1} \\ &= \sum_{k=1}^{m+n} \sum_{i+j=k} (i+j) a_i b_j t^{i+j-1} \\ &= \sum_{k=1}^{m+n} \sum_{i+j=k} i a_i t^{i-1} b_j t^j + \sum_{k=1}^{m+n} \sum_{i+j=k} a_i t^i j b_j t^{j-1} \\ &= \sum_{i=1}^{m}\sum_{j=0}^{n} i a_i t^{i-1} b_j t^j + \sum_{i=0}^{m}\sum_{j=1}^{n} a_i t^i j b_j t^{j-1} \\ &= D(f)g + fD(g). \end{aligned}$$

Therefore, D is a derivation on $R[t]$. □

An *integral domain* is a ring R such that if $b_1, b_2 \in R$ with $b_1 \neq 0$ and $b_2 \neq 0$, then $b_1 b_2 \neq 0$. Corresponding to every integral domain is a field, called the *quotient field* of R. It consists of all fractions of the form a/b,

where $a, b \in R$ and $b \neq 0$, and $a_1/b_1 = a_2/b_2$ if and only if $a_1b_2 = a_2b_1$. Addition and multiplication of fractions are defined in the usual way: If $a_1, a_2, b_1, b_2 \in R$ with $b_1 \neq 0$ and $b_2 \neq 0$, then $b_1b_2 \neq 0$ and

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1b_2 + a_2b_1}{b_1b_2} \qquad \text{and} \qquad \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1a_2}{b_1b_2}.$$

The quotient field of $\mathbf{Z}$ is $\mathbf{Q}$. If $F[t]$ is the ring of polynomials with coefficients in a field F, then the quotient field of $F[t]$ is the field $F(t)$ of rational functions with coefficients in F. A careful construction of quotient fields can be found in the Exercises.

Theorem 5.6 *Let R be an integral domain with quotient field F, and let D be a derivation on R. There exists a unique derivation D_F on F such that $D_F(x) = D(x)$ for all $x \in R$.*

Proof. Suppose that there exists a derivation D_F on F such that $D_F(a) = D(a)$ for all $a \in R$. Let $x \in F, x \neq 0$. There exist $a, b \in R$ with $b \neq 0$ and $x = a/b$. Since $a = bx \in R$, it follows that

$$D(a) = D_F(a) = D_F(bx) = D_F(b)x + bD_F(x) = D(b)x + bD_F(x),$$

and so

$$D_F\left(\frac{a}{b}\right) = D_F(x) = \frac{D(a) - D(b)x}{b} = \frac{D(a)b - aD(b)}{b^2}. \tag{5.3}$$

Thus, the derivation D_F on F is uniquely determined by the derivation D on R. In Exercise 3 we prove that (5.3) defines a derivation on the quotient field R_F. □

Let D be a derivation on the field F. For $x \in F^\times$, we define the *logarithmic derivative* $L(x)$ by

$$L(x) = \frac{D(x)}{x}.$$

If $x, y \in F^\times$, then

$$L(xy) = \frac{D(xy)}{xy} = \frac{D(x)y + xD(y)}{xy} = \frac{D(x)}{x} + \frac{D(y)}{y} = L(x) + L(y)$$

and

$$L\left(\frac{x}{y}\right) = \frac{D(x)}{x} + \frac{D(y^{-1})}{y^{-1}} = \frac{D(x)}{x} - \frac{D(y)}{y} = L(x) - L(y)$$

by Exercise 1.

We now consider polynomials with complex coefficients. A field F is called *algebraically closed* if every nonconstant polynomial with coefficients

in F has at least one zero in F. By the fundamental theorem of algebra, the field $\mathbf{C}$ is algebraically closed. Let $f(t) \in \mathbf{C}[t]$, and let $N_0(f)$ denote the number of distinct zeros of $f(t)$. If $f(t)$ has degree n with leading coefficient a_n, then $f(t)$ factors uniquely in the form

$$f(t) = a_n \prod_{i=1}^{N_0(f)} (t - \alpha_i)^{n_i},$$

where $\alpha_1, \ldots, \alpha_{N_0(f)}$ are the distinct zeros of f, the positive integer n_i is the multiplicity of the zero α_i, and $n_1 + \cdots + n_{N_0(f)} = n$. If D is the derivation on $\mathbf{C}[t]$ defined in Theorem 5.5, then, by Exercise 2,

$$D(f) = a_n \sum_{i=1}^{N_0(f)} n_i (t - \alpha_i)^{n_i - 1} \prod_{\substack{j=1 \\ j \neq i}}^{N_0(f)} (t - \alpha_j)^{n_j}$$

and

$$L(f) = \frac{D(f)}{f} = \sum_{i=1}^{N_0(f)} \frac{n_i}{t - \alpha_i}.$$

Let $g(t) = b_m \prod_{j=1}^{N_0(g)} (t - \beta_j)^{m_j}$ be a nonzero polynomial in $\mathbf{C}[t]$, and consider the rational function $f/g \in \mathbf{C}(t)$. Then

$$L\left(\frac{f}{g}\right) = L(f) - L(g) = \sum_{i=1}^{N_0(f)} \frac{n_i}{t - \alpha_i} - \sum_{j=1}^{N_0(g)} \frac{m_j}{t - \beta_j}. \tag{5.4}$$

This algebraic identity will be used in the next section to prove Mason's theorem.

Exercises

1. Let D be a derivation on a ring R. Prove that $D(1) = 0$ and that, if $x \in R$ is invertible, then

$$D(x^{-1}) = -\frac{D(x)}{x^2}.$$

2. Let D be a derivation on the ring R. Prove that

$$D(x_1 \cdots x_n) = \sum_{i=1}^{n} x_1 \cdots x_{i-1} D(x_i) x_{i+1} \cdots x_n$$

for all $x_1, \ldots, x_n \in R$.

3. Let R be an integral domain with quotient field F. Let D be a derivation on R, and define the function D_F on F by (5.3). We shall prove that D_F is a derivation on the quotient field F.

(a) Prove that D_F is well defined, that is, if $a_1/b_1 = a_2/b_2$, then $D_F(a_1/b_1) = D_F(a_2/b_2)$.

(b) Prove that

$$D_F\left(\frac{a_1}{b_1} + \frac{a_2}{b_2}\right) = D_F\left(\frac{a_1}{b_1}\right) + D_F\left(\frac{a_2}{b_2}\right).$$

(c) Prove that

$$D_F\left(\frac{a_1}{b_1} \cdot \frac{a_2}{b_2}\right) = D_F\left(\frac{a_1}{b_1}\right)\frac{a_2}{b_2} + \frac{a_1}{b_1}D_F\left(\frac{a_2}{b_2}\right).$$

4. Let R be a commutative ring with identity. A *multiplicatively closed* subset of R is a subset S such that $1 \in S$ and if $s_1, s_2 \in S$, then $s_1s_2 \in S$. We consider the set of ordered pairs of the form (r, s) with $r \in R$ and $s \in S$. Define a relation on this set as follows:

$$(r, s) \sim (r', s') \quad \text{if} \quad s''(s'r - sr') = 0 \text{ for some } s'' \in S.$$

Prove that this is an equivalence relation.

5. Let $S^{-1}R$ be the set of equivalence classes of the relation defined in Exercise 4. We denote the equivalence class of (r, s) by the fraction r/s. We also denote the equivalence class $(r, 1)$ by r. Define multiplication of fractions as follows:

$$\frac{r_1}{s_1} \cdot \frac{r_2}{s_2} = \frac{r_1r_2}{s_1s_2}.$$

(a) Prove that this multiplication is well defined, that is, if $(r_1, s_1) \sim (r_1', s_1')$ and $(r_2, s_2) \sim (r_2', s_2')$, then $(r_1r_2, s_1s_2) \sim (r_1'r_2', s_1's_2')$.

(b) Prove that multiplication in $S^{-1}R$ is associative and commutative, and that the equivalence class of $(1, 1)$ is a multiplicative identity.

(c) Prove that the equivalence class of $(s, 1)$ is invertible in $S^{-1}R$ for every $s \in S$.

(d) Prove that

$$\frac{a}{s} = \frac{s'a}{s's}$$

for all $a \in R$ and $s, s' \in S$.

6. Define addition of fractions in $S^{-1}R$ as follows:

$$\frac{r_1}{s_1} + \frac{r_2}{s_2} = \frac{s_2r_1 + s_1r_2}{s_1s_2}.$$

(a) Prove that this addition is well defined, that is, if $(r_1, s_1) \sim (r_1', s_1')$ and $(r_2, s_2) \sim (r_2', s_2')$, then $(s_2r_1 + s_1r_2, s_1s_2) \sim (s_2'r_1' + s_1'r_2', s_1's_2')$.

(b) Prove that addition in $S^{-1}R$ is associative and commutative, and that multiplication distributes over addition. Prove that the equivalence class of $(0, 1)$ is an additive identity.

7. (Localization) In Exercises 4–6 we proved that $S^{-1}R$ is a ring. This ring is called the *ring of fractions of R by S*. We also say that $S^{-1}R$ is constructed by *localizing* R at S.

(a) Prove that if $0 \in S$, then $S^{-1}R = \{0\}$.

(b) Prove that if R is an integral domain and $0 \notin S$, then $S^{-1}R$ is an integral domain.

(c) Prove that if R is an integral domain and S is the set of all nonzero elements of R, then $S^{-1}R$ is a field. This field is called the *quotient field* of the integral domain R.

8. Define $\varphi_S : R \to S^{-1}R$ by $\varphi_S(r) = r/1 = r$.

(a) Prove that φ_S is a ring homomorphism.

(b) Prove that if R is an integral domain and $0 \notin S$, then φ_R is one-to-one.

(c) Prove that if R is an integral domain and $S = R^\times$, then $S^{-1}R$ is isomorphic to R.

Hint: If S is a multiplicative subset of R and $s \in S \cap R^\times$, then $(r, s) \sim (s^{-1}r, 1)$ for all $r \in R$.

9. Let $S = \{1, 2, 4, 8, \ldots\}$ be the multiplicative subset of $\mathbf{Z}$ consisting of the powers of 2. Describe the ring of fractions $S^{-1}\mathbf{Z}$. What is the group of units in this ring?

10. Let $S = \{\pm 1, \pm 3, \pm 5, \pm 7, \ldots\}$ be the multiplicative subset of $\mathbf{Z}$ consisting of the odd integers.

(a) Describe the ring of fractions $S^{-1}\mathbf{Z}$.

(b) Describe the principal ideal generated by 2 in this ring.

(c) Prove that every element of the ring not in this ideal is a unit in $S^{-1}\mathbf{Z}$, and so $\langle 2 \rangle$ is a maximal ideal in $S^{-1}\mathbf{Z}$.

11. Let p be a prime number and let S be the set of all integers not divisible by p. Prove that S is a multiplicative subset of $\mathbf{Z}$, and describe the ring of fractions $S^{-1}\mathbf{Z}$. Prove that the principal ideal generated by p is a maximal ideal in $S^{-1}\mathbf{Z}$.

12. Let $F[t]$ be the polynomial ring with coefficients in the field F. Let $S = \{1, t, t^2, t^3, \ldots\}$ be the multiplicative subset of $F[t]$ consisting of the powers of t. Prove that $S^{-1}F[t]$ is isomorphic to the ring of *Laurent polynomials* with coefficients in F, that is, the ring consisting of all expressions of the form $\sum_{i=m}^{n} a_i t^i$, where $a_i \in F$, and m and n are integers with $m \leq n$, and addition and multiplication are defined in the usual way.

13. We consider the ring $R = \mathbf{Z}/12\mathbf{Z}$, and denote the congruence class $a + 12\mathbf{Z}$ by $\overline{a}$

    (a) Prove that $S = \{\overline{1}, \overline{3}, \overline{9}\}$ is a multiplicative subset of R.

    (b) Let $\varphi_S : R \to S^{-1}R$ be the ring homomorphism constructed in Exercise 8. Prove that $\varphi_S(\overline{a}) = \varphi_S(\overline{b})$ if and only if $a \equiv b \pmod{4}$.

    (c) Prove that $\overline{1}/\overline{3} = \overline{3}$ in $S^{-1}R$.

    (d) Prove that $S^{-1}R \cong \mathbf{Z}/4\mathbf{Z}$.

14. Let $m \geq 2$. We consider the ring $R = \mathbf{Z}/m\mathbf{Z}$, and denote the congruence class $a + m\mathbf{Z}$ by $\overline{a}$. Let S be a multiplicative subset of R such that $0 \notin S$.

    (a) Prove that we can factor m uniquely in the form $m = m_0 m_1$, where $(m_0, m_1) = 1$, and if p is a prime number that divides m, then p divides m_0 if and only if there is a congruence class $\overline{s} \in S$ such that p divides s. Show that $(s, m_1) = 1$ for all $\overline{s} \in S$.

    (b) Prove that there is a congruence class $\overline{s_0} \in S$ such that m_0 divides s_0.

    (c) Let $\varphi_S : R \to S^{-1}R$ be the ring homomorphism constructed in Exercise 8. Prove that $\varphi_S(\overline{a}) = \varphi_S(\overline{b})$ if and only if $a \equiv b \pmod{m_1}$.

    (d) Prove that for every $\overline{s} \in S$ there exists $\overline{r} \in R$ such that $\overline{1}/\overline{s} = \overline{r}$ in $S^{-1}R$.

    Hint: If $\overline{s} \in S$, then there exists an integer r such that $rs \equiv 1 \pmod{m_1}$.

    (e) Prove that $S^{-1}R \cong \mathbf{Z}/m_1\mathbf{Z}$.

5.3 Mason's Theorem

This is an important diophantine inequality for polynomials.

Theorem 5.7 (Mason) *If $a, b, c \in \mathbf{C}[t]$ are nonzero, relatively prime polynomials, not all constant, and if*

$$a + b = c,$$

then

$$\max\{\deg(a), \deg(b), \deg(c)\} \leq N_0(abc) - 1 = \deg(rad(abc)) - 1,$$

where $N_0(abc)$ denotes the number of distinct zeros of the polynomial abc, and rad(abc) is the radical of abc.

Since Mason's theorem is symmetric in a, b, and c, we could also write the equation in the form $a + b + c = 0$.

Proof. Let D be the unique derivation defined on the rational function field $\mathbf{C}(t)$ by Theorems 5.5 and 5.6, and let L be the logarithmic derivative. We introduce the nonzero rational functions $u = a/c$ and $v = b/c$ in $\mathbf{C}(t)$. Then $u + v = 1$, and

$$\begin{aligned} uL(u) + vL(v) &= u\left(\frac{D(u)}{u}\right) + v\left(\frac{D(v)}{v}\right) \\ &= D(u) + D(v) = D(u + v) = D(1) \\ &= 0. \end{aligned}$$

Since $L(v) \neq 0$ (by Exercise 1), we have

$$\frac{b}{a} = \frac{v}{u} = -\frac{L(u)}{L(v)}. \tag{5.5}$$

We write the standard factorizations of the polynomials a, b, and c as follows:

$$\begin{aligned} a &= a(t) = a_n \prod_{i=1}^{N_0(a)} (t - \alpha_i)^{n_i}, \\ b &= b(t) = b_m \prod_{i=1}^{N_0(b)} (t - \beta_i)^{m_i}, \\ c &= c(t) = c_r \prod_{i=1}^{N_0(c)} (t - \gamma_i)^{r_i}. \end{aligned}$$

Applying (5.4), we obtain

$$L(u) = L\left(\frac{a}{c}\right) = \sum_{i=1}^{N_0(a)} \frac{n_i}{t - \alpha_i} - \sum_{j=1}^{N_0(c)} \frac{r_k}{t - \gamma_k}$$

and

$$L(v) = L\left(\frac{b}{c}\right) = \sum_{j=1}^{N_0(b)} \frac{m_j}{t-\beta_j} - \sum_{j=1}^{N_0(c)} \frac{r_k}{t-\gamma_k}.$$

Since the polynomials a, b, and c are relatively prime, the radical of the product abc is

$$q = \mathrm{rad}(abc) = \prod_{i=1}^{N_0(a)} (t-\alpha_i) \prod_{i=1}^{N_0(b)} (t-\beta_i) \prod_{i=1}^{N_0(c)} (t-\gamma_i),$$

and

$$\deg(q) = \deg(\mathrm{rad}(abc)) = N_0(a) + N_0(b) + N_0(c).$$

Moreover, $qL(u)$ and $qL(v)$ are polynomials of degree at most $\deg(q) - 1$. By (5.5),

$$\frac{b}{a} = -\frac{L(u)}{L(v)} = -\frac{qL(u)}{qL(v)},$$

and so

$$a(qL(u)) = -b(qL(v)).$$

Since the polynomials a and b are relatively prime, it follows that a divides $qL(v)$, and so

$$\deg(a) \leq \deg(qL(v)) \leq \deg(q) - 1 = \deg(\mathrm{rad}(abc)) - 1.$$

Similarly,

$$\deg(b) \leq \deg(qL(u)) \leq \deg(q) - 1 = \deg(\mathrm{rad}(abc)) - 1$$

and

$$\deg(c) \leq \deg(\mathrm{rad}(abc)) - 1.$$

This completes the proof. □

Fermat's last theorem states that if $n \geq 3$, then the *Fermat equation*

$$x^n + y^n = z^n$$

has no solutions in positive integers. The Fermat equation has solutions in polynomials for $n = 2$, for example,

$$(1-t^2)^2 + (2t)^2 = (1+t^2)^2.$$

We shall use Mason's theorem to prove Fermat's last theorem for polynomials for $n \geq 3$.

Theorem 5.8 *If $n \geq 3$, then the Fermat equation $x^n + y^n = z^n$ has no solution in nonzero, relatively prime polynomials, not all constant.*

Proof. Let $n \geq 3$, and suppose that x, y, and z are nonzero, relatively prime polynomials, not all constant, such that $x^n + y^n = z^n$. We apply Mason's theorem with $a = x^n, b = y^n$, and $c = z^n$. Then

$$\mathrm{rad}(abc) = \mathrm{rad}(x^n y^n z^n) = \mathrm{rad}(xyz).$$

Since $\deg(x^n) = n\deg(x)$, we obtain

$$\begin{aligned}
n\deg(x) &\leq n\max(\deg(x), \deg(y), \deg(z)) \\
&= \max(\deg(x^n), \deg(y^n), \deg(z^n)) \\
&= \max(\deg(a), \deg(b), \deg(c)) \\
&\leq \deg(\mathrm{rad}(abc)) - 1 \\
&= \deg(\mathrm{rad}(xyz)) - 1 \\
&\leq \deg(xyz) - 1 \\
&= \deg(x) + \deg(y) + \deg(z) - 1.
\end{aligned}$$

It follows that

$$\begin{aligned}
n(\deg(x) + \deg y + \deg(z)) &\leq 3(\deg(x) + \deg y + \deg(z)) - 3 \\
&\leq n(\deg(x) + \deg y + \deg(z)) - 3.
\end{aligned}$$

This is impossible. □

Exercises

1. Prove that $L(v) \neq 0$ in the proof of Theorem 5.7.

2. Let $n \geq 3$. Prove that the equation $x^n + y^n = 1$ has no solution in nonconstant rational functions $x, y \in \mathbf{C}(t)$.

3. (Nathanson [102]) The *Catalan equation* is the equation

$$x^m - y^n = 1,$$

where m and n are integers greater than 1. Prove that this equation has no solution in nonconstant polynomials $x, y \in \mathbf{C}[t]$ and integers $m \geq 2$ and $n \geq 2$.

4. (Davenport [20]) Let f and g be nonconstant, relatively prime polynomials in $\mathbf{C}[t]$. Prove that

$$\deg(f^3 - g^2) \geq \frac{1}{2}\deg(f) + 1.$$

5. Let

$$\begin{aligned} f &= t^6 + 4t^4 + 10t^2 + 6 \\ g &= t^9 + 6t^7 + 21t^5 + 35t^3 + \frac{63}{2}t. \end{aligned}$$

Check that

$$f^3 - g^2 = 27t^4 + \frac{351}{4}t^2 + 216.$$

This example shows that the lower bound in Davenport's theorem (Exercise 4) is best possible.

5.4 The *abc* Conjecture

The *abc* conjecture is a simple but powerful assertion about the relationship between the additive and multiplicative properties of integers. Recall that the radical of a nonzero integer m is the largest square-free divisor of m, that is,

$$\mathrm{rad}(m) = \prod_{p|m} p.$$

The *abc conjecture* states that for every $\varepsilon > 0$ there exists a number $K(\varepsilon)$ such that, if a, b, and c are nonzero, relatively prime integers and

$$a + b = c,$$

then

$$\max(|a|, |b|, |c|) \leq K(\varepsilon)\mathrm{rad}(abc)^{1+\varepsilon}.$$

Since the inequality is symmetric in a, b, and c, the equation can also be written in the form $a + b + c = 0$. To prove or disprove this conjecture is an important unsolved problem in number theory.

From the *abc* conjecture it is possible to deduce many theorems and still unproven propositions in number theory. Here are some examples.

Fermat's last theorem states that, for $n \geq 3$, the *Fermat equation*

$$x^n + y^n = z^n \tag{5.6}$$

has no solution in positive integers. Note that if x, y, z is a solution of (5.6) in positive integers and if a prime number p divides x and y, then p also divides z, and $x/p, y/p, z/p$ is another solution of the equation. It follows that if the Fermat equation has a solution in integers, then it has a solution in relatively prime integers.

Theorem 5.9 (Asymptotic Fermat theorem) *The abc conjecture implies that there exists an integer n_0 such that the Fermat equation has no solution in relatively prime integers for any exponent $n \geq n_0$.*

Proof. Let x, y, and z be relatively prime positive integers such that

$$x^n + y^n = z^n.$$

We note that

$$\text{rad}(x^n y^n z^n) = \text{rad}(xyz) \leq xyz \leq z^3.$$

If $n \geq 2$, then $z \geq 3$. Applying the *abc* conjecture with $\varepsilon = 1$ and $K_1 = \max(1, K(1))$, we obtain

$$z^n = \max(x^n, y^n, z^n) \leq K_1 \text{rad}(x^n y^n z^n)^2 < K_1 z^6,$$

and so

$$n < 6 + \frac{\log K_1}{\log z} \leq 6 + \frac{\log K_1}{\log 3}.$$

This completes the proof. □

The *Catalan conjecture* asserts that 8 and 9 are the only consecutive powers. Equivalently, it states that the only solution of the *Catalan equation*

$$x^m - y^n = 1$$

in integers x, y, m, n all greater than 1 is

$$3^2 - 2^3 = 1.$$

It is known that the diophantine equation $x^m - y^2 = 1$ has no solution in positive integers, and that the only solution of the equation $x^2 - y^n = 1$ in positive integers is $x = n = 3$ and $y = 2$. Therefore, it suffices to consider the Catalan equation only for $\min(m, n) \geq 3$.

Theorem 5.10 (Asymptotic Catalan theorem) *The abc conjecture implies that the Catalan equation has only finitely many solutions.*

Proof. Let (x, y, m, n) be a solution of the Catalan equation with $\min(m, n) \geq 3$. Then x and y are relatively prime. It follows from the *abc* conjecture with $\varepsilon = 1/4$ that there exists a constant $K_2 = K(1/4)$ such that

$$y^n < x^m \leq K_2 \text{rad}(x^m y^n)^{5/4} = K_2 \text{rad}(xy)^{5/4} \leq K_2 (xy)^{5/4},$$

and so

$$m \log x \leq \log K_2 + \frac{5}{4} (\log x + \log y)$$

and

$$n \log y < \log K_2 + \frac{5}{4} (\log x + \log y).$$

It follows that

$$m \log x + n \log y < 2 \log K_2 + \frac{5}{2} (\log x + \log y),$$

and so

$$\left(m - \frac{5}{2}\right) \log x + \left(n - \frac{5}{2}\right) \log y < 2 \log K_2. \tag{5.7}$$

Since $x \geq 2$ and $y \geq 2$, we have

$$m + n < \frac{2 \log K_2}{\log 2} + 5.$$

Thus, there are only finitely many pairs of exponents (m, n) for which the Catalan equation is solvable. For fixed exponents $m \geq 3$ and $n \geq 3$, equation (5.7) has only only finitely many solutions in positive integers x and y. This completes the proof. □

For every odd prime p we have $2^{p-1} \equiv 1 \pmod{p}$, that is, p divides $2^{p-1} - 1$. The question of the divisibility of $2^{p-1} - 1$ by p^2 arose in the study of Fermat's last theorem. An odd prime p such that

$$2^{p-1} \not\equiv 1 \pmod{p^2}$$

is called a *Wieferich prime*. For example, 3, 5, and 7 are Wieferich primes, since $2^2 \not\equiv 1 \pmod{9}$, $2^4 \not\equiv 1 \pmod{25}$, and $2^6 \not\equiv 1 \pmod{49}$. It is not known whether infinitely many Wieferich primes exist, nor is is known whether there are infinitely many primes that are not Wieferich primes.

Let W be the set of Wieferich primes. We shall show that the *abc* conjecture implies that W is infinite. We begin with a simple lemma.

Lemma 5.1 *Let p be an odd prime. If there exists a positive integer n such that $2^n \equiv 1 \pmod{p}$ but $2^n \not\equiv 1 \pmod{p^2}$, then p is a Wieferich prime.*

Proof. Let d be the order of 2 modulo p. Then d divides n. Since $2^n \not\equiv 1 \pmod{p^2}$, it follows that $2^d \not\equiv 1 \pmod{p^2}$. Then $2^d = 1 + kp$, where $(k, p) = 1$. Moreover, d divides $p - 1$, since $2^{p-1} \equiv 1 \pmod{p}$, and so $p - 1 = de$ for some integer e such that $1 \leq e \leq p - 1$. Then $(ek, p) = 1$ and

$$2^{p-1} = (2^d)^e = (1 + kp)^e \equiv 1 + ekp \not\equiv 1 \pmod{p^2},$$

and p is a Wieferich prime. □

A *powerful number* is a positive integer v such that if a prime p divides v, then p^2 divides v. For example, 72 is powerful but 192 is not. If v is powerful, then $\text{rad}(v) \leq v^{1/2}$.

Theorem 5.11 *The abc conjecture implies that there exist infinitely many Wieferich primes.*

Proof. Let W be the set of Wieferich primes. For every positive integer n, we write

$$2^n - 1 = u_n v_n,$$

where v_n is the maximal powerful divisor of $2^n - 1$. Then u_n is a square-free integer,

$$u_n = \prod_{\substack{p|n \\ v_p(n)=1}} p,$$

and

$$v_n = \prod_{\substack{p|n \\ v_p(n)\geq 2}} p^{v_p(n)}.$$

If p divides u_n, then

$$2^n \equiv 1 \pmod{p}$$

but

$$2^n \not\equiv 1 \pmod{p^2}.$$

It follows from Lemma 5.1 that $p \in W$, and so u_n is a square-free integer divisible only by Wieferich primes.

If the set W is finite, then there exist only finitely many square-free integers whose prime divisors all belong to W, and so the set $\{u_n : n = 1, 2, 3, \ldots\}$ is finite. It follows that the set $\{v_n : n = 1, 2, 3, \ldots\}$ is infinite, and, consequently, unbounded. Since v_n is powerful, we have

$$\mathrm{rad}(v_n) \leq v_n^{1/2}.$$

Let $0 < \varepsilon < 1$. Applying the *abc* conjecture to the identity

$$(2^n - 1) + 1 = 2^n,$$

we obtain

$$\begin{aligned} v_n &< 2^n \\ &\leq K(\varepsilon)\mathrm{rad}(2^n(2^n-1))^{1+\varepsilon} \\ &\leq K(\varepsilon)\mathrm{rad}(2u_n v_n)^{1+\varepsilon} \\ &\leq K(\varepsilon)(2u_n)^{1+\varepsilon}\mathrm{rad}(v_n)^{1+\varepsilon} \\ &\ll v_n^{(1+\varepsilon)/2}. \end{aligned}$$

This implies that the numbers v_n are bounded, which is absurd. This completes the proof. □

Exercises

1. For a fixed exponent $n \geq 4$, prove that the Fermat equation $x^n + y^n = z^n$ has at most a finite number of solutions in positive integers x, y, z. Does this argument show that the cubic Fermat equation $x^3 + y^3 = z^3$ has at most finitely many solutions?

   Hint: Apply the *abc* conjecture with $\varepsilon = 1/6$.

2. An integer n is *powerful* if $v_p(n) \neq 1$ for all primes p. Compute the powerful numbers up to 100.

3. Let $n \geq 2$ be an integer. Define the *power of n* by

$$\text{power}(n) = \frac{\log n}{\log \text{rad}(n)}.$$

   Prove that $\text{power}(n) = 1$ if and only if n is square-free. Prove that if n is powerful, then $\text{power}(n) \geq 2$. Prove that if n is a kth power, then $\text{power}(n) \geq k$.

4. (Granville) Prove that the *abc* conjecture implies that there exist only finitely many triples of consecutive powerful numbers.

   Hint: Suppose that $n-1, n, n+1$ are three consecutive powerful numbers. Apply the *abc* conjecture to the equation $(n^2-1)+1 = n^2$. Observe that

$$\begin{aligned} \text{rad}(n^2(n^2-1)) &= \text{rad}((n-1)n(n+1)) \\ &\leq \sqrt{(n-1)n(n+1)} < n^{3/2}. \end{aligned}$$

5. Let

$$U = \bigcup_{k=3}^{\infty} \{x^k : x \in \mathbf{N}\} = \{u_i\}_{i=1}^{\infty}$$

   be the set of nonsquare powers of the positive integers, where $u_i < u_{i+1}$ for $i = 1, 2, \ldots$. Prove that the *abc* conjecture implies

$$\lim_{i \to \infty} (u_{i+1} - u_i) = \infty.$$

6. Prove that the *abc* conjecture implies that the diophantine equation

$$n! + 1 = m^2$$

   has only finitely many solutions.

   Hint: Apply the inequalities

$$\prod_{p \leq n} p < 4^n$$

(Theorem 8.1) and

$$\frac{1}{n}\left(\frac{n}{e}\right)^n < n! < \left(\frac{n}{e}\right)^n$$

(Exercise 1 in Section 6.2).

7. Prove that the *abc* conjecture is false if we omit the condition $(a, b, c) = 1$.

   Hint: Consider the equation $3^k + 2 \cdot 3^k = 3^{k+1}$.

8. In this exercise we construct an example to show that the *abc* conjecture would be false if we replaced the exponent $1 + \varepsilon$ with 1.

   (a) Prove that for every positive integer n there exists a positive integer u_n such that

   $$2^n u_n + 1 = 3^{2^{n-1}}.$$

   Hint: Euler's theorem.

   (b) Let $a_n = 2^n u_n$, $b_n = 1$, and $c_n = 3^{2^{n-1}}$. Prove that

   $$\mathrm{rad}(a_n b_n c_n) = \mathrm{rad}\,(6u_n) < \frac{6 \cdot 3^{2^{n-1}}}{2^n}.$$

   (c) Let $K(0) > 0$. Prove that if n is sufficiently large, then

   $$K(0)\mathrm{rad}(a_n b_n c_n) < \frac{6K(0)c_n}{2^n} < c_n = \max(a_n, b_n, c_n).$$

   Since $a_n + b_n = c_n$, this is the desired counterexample.

9. Let a and b be relatively prime positive integers. We define $c = a + b$ and

   $$L(a, b) = \frac{\log c}{\log \mathrm{rad}(abc)} = \frac{\log(a+b)}{\log \mathrm{rad}(ab(a+b))}.$$

   It is hard to find relatively prime integers a and b for which $L(a, b)$ is large. Use the equation

   $$2 + 3^{10}109 = 23^5$$

   to compute $L(2, 3^{10}109)$. In October,1999, this was the largest known value for $L(a, b)$.

10. Compute $L(a, b)$ for $a = 1$ and $b = 2 \cdot 3^7$.

11. Compute $L(a, b)$ for $a = 11^2$ and $b = 3^2 \cdot 5^6 \cdot 7^3$.

12. For $n \geq 1$, define the positive integer t_n by

$$9^n = 1 + 8t_n.$$

Prove that $L(1, 8t_n) > 1$ and so

$$\limsup_{(a,b)=1} L(a,b) \geq 1.$$

It can be shown that the *abc* conjecture is equivalent to

$$\limsup_{(a,b)=1} L(a,b) = 1.$$

5.5 The Congruence *abc* Conjecture

Let $m \geq 2$. The *congruence abc conjecture for* m states that for every $\varepsilon > 0$ there exists a number $K(m,\varepsilon)$ such that, if a, b, c are nonzero, relatively prime integers with

$$abc \equiv 0 \pmod{m}$$

and

$$a + b = c,$$

then

$$\max(|a|, |b|, |c|) \leq K(m,\varepsilon)\mathrm{rad}(abc)^{1+\varepsilon}.$$

This a weaker assertion than the *abc* conjecture, which is unrestricted by any congruence condition. However, we shall prove that if the congruence *abc* conjecture is true for some modulus m, then the unrestricted *abc* conjecture is also true.

We begin with some simple observations about triples (a, b, c) of integers such that $a + b = c$. First, at least one of the integers a, b, or c must be even, and so $abc \equiv 0 \pmod{2}$. Therefore, the congruence *abc* conjecture for $m = 2$ is the same as the *abc* conjecture, and we need to consider only moduli $m \geq 3$. Second, if $(a, b, c) = 1$, then either c is odd and $b - a$ is odd, or c is even, both a and b are odd, and $b - a$ is even. Third, if a, b, c are distinct nonzero integers, then, by a permutation, we can assume that they are positive and $a < b < c$.

Lemma 5.2 *Let a, b, c be relatively prime positive integers such that*

$$a < b < c$$

and

$$a + b = c.$$

Let $n \geq 2$. If c is odd, define

$$\begin{aligned} A_n &= (b-a)^n, \\ B_n &= c^n - (b-a)^n, \\ C_n &= c^n. \end{aligned}$$

If c is even, define

$$\begin{aligned} A_n &= \left(\frac{b-a}{2}\right)^n, \\ B_n &= \left(\frac{c}{2}\right)^n - \left(\frac{b-a}{2}\right)^n, \\ C_n &= \left(\frac{c}{2}\right)^n. \end{aligned}$$

Then A_n, B_n, C_n are distinct, relatively prime positive integers such that

$$A_n + B_n = C_n.$$

If $m \geq 3$ and $n = \varphi(m)$, then

$$A_n B_n C_n \equiv 0 \pmod{m}.$$

Proof. It is left to the reader to show that A_n, B_n, C_n are distinct, relatively prime positive integers such that $A_n + B_n = C_n$ (Exercises 1, 2, and 3).

Let $m \geq 3$ and $n = \varphi(m)$. Then $n \geq 2$. We must prove that

$$A_n B_n C_n \equiv 0 \pmod{m}.$$

It suffices to prove that if p is a prime and p^r divides m, then

$$A_n B_n C_n \equiv 0 \pmod{p^r}. \tag{5.8}$$

Note that if p is a prime and p^r divides m, then $(p-1)p^{r-1}$ divides n, and so

$$r \leq 2^{r-1} \leq (p-1)p^{r-1} \leq n.$$

Suppose that p is an odd prime. If p divides c, then p^n divides c^n and p^n divides C_n. Since $r \leq n$, it follows that $C_n \equiv 0 \pmod{p^r}$. Similarly, if p divides $b-a$, then $A_n \equiv 0 \pmod{p^r}$. If p divides neither c nor $b-a$, then, by Theorem 2.12,

$$c^{(p-1)p^{r-1}} \equiv 1 \pmod{p^r}$$

and

$$(b-a)^{(p-1)p^{r-1}} \equiv 1 \pmod{p^r}.$$

Since $(p-1)p^{r-1}$ divides n, we have

$$c^n \equiv (b-a)^n \equiv 1 \pmod{p^r},$$

and so $B_n \equiv 0 \pmod{p^r}$. This proves (5.8) for odd primes p.

Finally, we consider the prime 2. If 2^r divides m, then 2^{r-1} divides n and $r \leq n$. If c is even, then $b-a$ is even and exactly one of the integers c and $b-a$ is divisible by 4 (Exercise 4). It follows that either c^n or $(b-a)^n$ is divisible by 4^n, and so either C_n or A_n is divisible by 2^n, which is divisible by 2^r.

If c is odd, then $b-a$ is odd and

$$c^{2^{r-1}} \equiv (b-a)^{2^{r-1}} \equiv 1 \pmod{2^r}.$$

Since 2^{r-1} divides n, we have

$$B_n = c^n - (b-a)^n \equiv 0 \pmod{2^r}.$$

This proves (5.8) for the prime 2. □

Theorem 5.12 *Let $m \geq 3$. If the congruence abc conjecture is true for m, then the abc conjecture is true.*

Proof. Let $0 < \varepsilon < 1$. For triples a, b, c of distinct, relatively prime positive integers such that $a + b = c$, we define the function

$$\Phi_\varepsilon(a, b, c) = \log c - (1+\varepsilon) \log \operatorname{rad}(abc).$$

Then

$$\log \operatorname{rad}(a, b, c) = \log c - \frac{\varepsilon \log c}{1+\varepsilon} - \frac{\Phi_\varepsilon(a, b, c)}{1+\varepsilon}.$$

Let A, B, C be distinct, relatively prime positive integers such that $ABC \equiv 0 \pmod{m}$ and $A + B = C$. If the congruence *abc* conjecture is true for m, then there exists a constant $K(m, \varepsilon) > 0$ such that

$$C \leq K(m, \varepsilon) \operatorname{rad}(ABC)^{1+\varepsilon},$$

or, equivalently,

$$\Phi_\varepsilon(A, B, C) \leq \log K(m, \varepsilon) = K^*(m, \varepsilon).$$

Let a, b, c be relatively prime positive integers such that $a < b < c$ and $a + b = c$. Let

$$n = \varphi(m).$$

Then n is even, by Exercise 4 in Section 2.3. Define the integers A_n, B_n, C_n as in Lemma 5.2. Then $A_n B_n C_n \equiv 0 \pmod{m}$ and $A_n + B_n = C_n$. Moreover,

$$\Phi_\varepsilon(A_n, B_n, C_n) \leq K^*(m, \varepsilon).$$

The integer n is even, since $m \geq 3$, and so, by Exercise 5,

$$\begin{aligned} B_n &= c^n - (b-a)^n \\ &= (b+a)^n - (b-a)^n \\ &= 4ab\left((b+a)^{n-2} + (b+a)^{n-4}(b-a)^2 + \cdots + (b-a)^{n-2}\right) \\ &\leq 4ab\left(\frac{n}{2}\right)(b+a)^{n-2} \\ &= 2abnc^{n-2}. \end{aligned}$$

Since

$$A_nB_nC_n = (b-a)^n\left(\frac{B_n}{ab}\right)abc^n,$$

it follows that

$$\begin{aligned} \operatorname{rad}(A_nB_nC_n) &= \operatorname{rad}\left((b-a)^n\left(\frac{B_n}{ab}\right)abc^n\right) \\ &= \operatorname{rad}\left((b-a)\left(\frac{B_n}{ab}\right)abc\right) \\ &\leq \operatorname{rad}(b-a)\operatorname{rad}\left(\frac{B_n}{ab}\right)\operatorname{rad}(abc) \\ &\leq (b-a)\left(\frac{B_n}{ab}\right)\operatorname{rad}(abc) \\ &\leq (b-a)\left(2nc^{n-2}\right)\operatorname{rad}(abc) \\ &\leq 2nc^{n-1}\operatorname{rad}(abc). \end{aligned}$$

Therefore,

$$\begin{aligned} \log\operatorname{rad}(A_nB_nC_n) &\leq (n-1)\log c + \log\operatorname{rad}(abc) + \log 2n \\ &= n\log c - \frac{\varepsilon\log c}{1+\varepsilon} - \frac{\Phi_\varepsilon(a,b,c)}{1+\varepsilon} + \log 2n \\ &= \left(1 - \frac{\varepsilon}{(1+\varepsilon)n}\right)\log c^n - \frac{\Phi_\varepsilon(a,b,c)}{1+\varepsilon} + \log 2n \\ &\leq \left(1 - \frac{\varepsilon}{(1+\varepsilon)n}\right)(\log C_n + n\log 2) - \frac{\Phi_\varepsilon(a,b,c)}{1+\varepsilon} + \log 2n \\ &\leq \left(\frac{n+(n-1)\varepsilon}{(1+\varepsilon)n}\right)\log C_n - \frac{\Phi_\varepsilon(a,b,c)}{1+\varepsilon} + 2n\log n. \end{aligned}$$

Equivalently,

$$\begin{aligned} \Phi_\varepsilon(a,b,c) &\leq \left(\frac{n+(n-1)\varepsilon}{n}\right)\left(\log C_n - \left(\frac{(1+\varepsilon)n}{n+(n-1)\varepsilon}\right)\log\operatorname{rad}(A_nB_nC_n)\right) \\ &\quad + 2(1+\varepsilon)n\log 2 \\ &< 2\left(\log C_n - \left(\frac{(1+\varepsilon)n}{n+(n-1)\varepsilon}\right)\log\operatorname{rad}(A_nB_nC_n)\right) + 4n\log 2 \\ &= 2\left(\log C_n - (1+\varepsilon')\log\operatorname{rad}(A_nB_nC_n)\right) + 4n\log 2, \end{aligned}$$

where

$$\varepsilon' = \frac{(1+\varepsilon)n}{n+(n-1)\varepsilon} - 1 = \frac{\varepsilon}{\varphi(m)+(\varphi(m)-1)\varepsilon}.$$

Since

$$\log C_n - (1+\varepsilon')\log \operatorname{rad}(A_nB_nC_n) = \Phi_{\varepsilon'}(A_n, B_n, C_n) \leq K^*(\varepsilon', m),$$

it follows that

$$\Phi_\varepsilon(a,b,c) < 2K^*(\varepsilon', m) + 4\varphi(m)\log 2.$$

Thus, for every $\varepsilon > 0$, the function $\Phi_\varepsilon(a,b,c)$ is bounded above, and this is equivalent to the *abc* conjecture. This completes the proof. □

Exercises

1. Let a, b, c positive integers such that $(a,b,c) = 1$ and $a+b = c$. Prove that $(a,b) = (a,c) = (b,c) = 1$. Prove that $a = b$ only if $a = 1$ and $c = 2$.

2. Let a, b, c be relatively prime positive integers such that c is odd, $a < b < c$, and
$$a + b = c.$$
For every positive integer n, define
$$\begin{aligned} A_n &= (b-a)^n, \\ B_n &= c^n - (b-a)^n, \\ C_n &= c^n. \end{aligned}$$
Prove that A_n, B_n, and C_n are distinct, relatively prime positive integers such that
$$A_n + B_n = C_n.$$

3. Let a, b, and c be relatively prime positive integers such that c is even, $a < b < c$, and
$$a + b = c.$$
For every positive integer n, define
$$\begin{aligned} A_n &= \left(\frac{b-a}{2}\right)^n, \\ B_n &= \left(\frac{c}{2}\right)^n - \left(\frac{b-a}{2}\right)^n, \\ C_n &= \left(\frac{c}{2}\right)^n. \end{aligned}$$

Prove that A_n, B_n, and C_n are distinct, relatively prime positive integers such that

$$A_n + B_n = C_n.$$

4. Let a, b, c be relatively prime integers such that $a + b = c$. Prove if c is even, then exactly one of the integers c and $b - a$ is divisible by 4.

5. Prove that if n is even, then

$$(b+a)^n-(b-a)^n = 4ab\left((b+a)^{n-2} + (b+a)^{n-4}(b-a)^2 + \cdots + (b-a)^{n-2}\right).$$

5.6 Notes

> One of the most fruitful analogies in mathematics is that between the integers $\mathbf{Z}$ and the ring of polynomials $F[t]$ over a field F.
>
> S. Lang [89, p. 196]

There are beautiful survey articles on the *abc* conjecture by Lang, "Old and new conjectured diophantine inequalities" [88], Nitaj, "La conjecture *abc*' [113], and Brzeziński, "The *abc*-conjecture" [15]. Part of Lang's article appears in his *Algebra* [89, pages 194–200], which is a highly recommended reference for all matters algebraical.

The *abc* conjecture was motivated in part by Mason's theorem, which is a polynomial analogue of the *abc* conjecture (see Mason [97]), and in part by a conjecture of Szpiro on the discriminants of elliptic curves (Lang [88]). According to Oesterlé [114, pp. 167–169], Szpiro had discussed this conjecture in a lecture in Hanover in 1983; the *abc* conjecture arose in a discussion between Masser and Oesterlé in 1985.

Browkin and Brzeziński [14] contains considerable data on the values of the function $L(a, b)$, discussed in Exercises (9)–(12), as well as a conjectured generalization of the *abc* conjecture to equations of the form $a_1 + a_2 + \cdots + a_n = 0$. The proof that the congruence *abc* conjecture implies the *abc* conjecture is due to Ellenberg [27].

Fermat's last theorem was proved by Taylor and Wiles [139, 156] in 1995. For a different proof of Fermat's last theorem for polynomials, see Greenleaf [41]. For a proof that the Catalan equation has no solution in polynomials or rational functions, see Nathanson [102].

V. A. Lebesgue [91] proved that the diophantine equation $x^m = y^2 + 1$ has no solution in positive integers. Chao Ko [82] proved that the only solution of $x^2 = y^m + 1$ in positive integers is $x = m = 3$ and $y = 2$.

Silverman [134] applied the *abc* conjecture to Wieferich primes (Theorem 5.11). Wieferich [155] proved that if p is an odd prime such that the

Fermat equation

$$x^p + y^p = z^p$$

has a solution in integers x, y, z with $(p, xyz) = 1$, then

$$2^{p-1} \equiv 1 \pmod{p^2}.$$

Computations [17] suggest that such primes are rare, and that "most" primes are Wieferich primes. Indeed, 1093 and 3511 are the only primes $p \leq 4 \cdot 10^{22}$ that are not Wieferich primes. It is an open problem to determine whether there exists a prime p that satisfies the following two congruences:

$$2^{p-1} \equiv 1 \pmod{p^2}$$

and

$$3^{p-1} \equiv 1 \pmod{p^2}.$$

Part II

Divisors and Primes in Multiplicative Number Theory

6

Arithmetic Functions

6.1 The Ring of Arithmetic Functions

An *arithmetic function* is a complex-valued function whose domain is the set of positive integers. For example, the divisor function $d(n)$ and the Euler phi function $\varphi(n)$ are arithmetic functions.

The *pointwise sum* $f+g$ of the arithmetic functions f and g is defined by

$$(f+g)(n) = f(n) + g(n). \tag{6.1}$$

There are two natural ways to multiply arithmetic functions f and g. The first is the *pointwise product* $f \cdot g$, defined by

$$f \cdot g(n) = f(n)g(n).$$

The second is the *Dirichlet convolution* $f * g$, defined by

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d) = \sum_{dd'=n} f(d)g(d'), \tag{6.2}$$

where the sum is over all positive divisors d of n. Dirichlet convolution occurs frequently in multiplicative problems in elementary number theory.

We define the arithmetic function $\delta(n)$ by

$$\delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \geq 2, \end{cases}$$

and the *zero function* $0(n)$ by $0(n) = 0$ for all n.

Theorem 6.1 *The set of all complex-valued arithmetic functions, with addition defined by pointwise sum and multiplication defined by Dirichlet convolution, is a commutative ring with additive identity* $0(n)$ *and multiplicative identity* $\delta(n)$.

Proof. It is easy to check that the set of arithmetic functions is an additive abelian group with the zero function as the additive identity.

We shall prove that Dirichlet convolution is commutative, associative, and distributes over addition, that is,

$$f * g = g * f,$$

$$(f * g) * h = f * (g * h),$$

and

$$f * (g + h) = f * g + f * h$$

for all arithmetic functions f, g, and h. These are straightforward calculations. We have

$$f * g(n) = \sum_{d|n} f(d)g(n/d) = \sum_{d|n} g(n/d)f(d) = \sum_{d|n} g(d)f(n/d) = g * f(n)$$

and

$$\begin{aligned}
((f * g) * h)(n) &= \sum_{d|n} (f * g)(d)h\left(\frac{n}{d}\right) \\
&= \sum_{dm=n} (f * g)(d)h(m) \\
&= \sum_{dm=n} \sum_{k\ell=d} f(k)g(\ell)h(m) \\
&= \sum_{k\ell m=n} f(k)g(\ell)h(m) \\
&= \sum_{k|n} f(k) \sum_{\ell m=n/k} g(\ell)h(m) \\
&= \sum_{k|n} f(k) \sum_{\ell|(n/k)} g(\ell)h\left(\frac{n}{k\ell}\right) \\
&= \sum_{k|n} f(k)(g * h)\left(\frac{n}{k}\right) \\
&= (f * (g * h))(n).
\end{aligned}$$

Similarly,

$$f * (g + h)(n) = \sum_{d|n} f(d)(g(n/d) + h(n/d))$$

$$
\begin{aligned}
&= \sum_{d|n} f(d)g(n/d) + \sum_{d|n} f(d)h(n/d) \\
&= f * g(n) + f * h(n).
\end{aligned}
$$

Finally, we observe that

$$\delta * f(n) = \sum_{d|n} \delta(d) f(n/d) = f(n)$$

for every arithmetic function f, and so the arithmetic functions form a commutative ring with multiplicative identity $\delta(n)$. This completes the proof. □

Recall that a *derivation* on a ring R is an additive homomorphism $D : R \to R$ such that

$$D(xy) = D(x)y + xD(y)$$

for all $x, y \in R$.

Theorem 6.2 *Consider the arithmetic function $L(n)$ defined by*

$$L(n) = \log n \quad \text{for all } n \geq 1.$$

Pointwise multiplication by $L(n)$ is a derivation on the ring of arithmetic functions.

Proof. Observe that if d is a positive divisor of n, then

$$L(n) = L(d) + L(n/d).$$

We must prove that

$$L \cdot (f * g) = (L \cdot f) * g + f * (L \cdot g)$$

for all arithmetic functions f and g. We have

$$
\begin{aligned}
L \cdot (f * g)(n) &= L(n) \sum_{d|n} f(d)g(n/d) \\
&= \sum_{d|n} L(n) f(d) g(n/d) \\
&= \sum_{d|n} (L(d) + L(n/d)) f(d) g(n/d) \\
&= \sum_{d|n} L(d) f(d) g(n/d) + \sum_{d|n} f(d) L(n/d) g(n/d) \\
&= (L \cdot f) * g + f * (L \cdot g).
\end{aligned}
$$

This completes the proof. □

Exercises

1. Define the arithmetic function $1(n)$ by $1(n) = 1$ for all n. Prove that $1 * 1(n) = d(n)$.

2. For every positive integer k, let $d_k(n)$ denote the number of k-tuples of positive integers $(a_1, a_2, \ldots, a_k)$ such that $n = a_1 a_2 \cdots a_k$. Prove that
$$d_k(n) = \underbrace{1 * 1 * \cdots * 1}_{k \text{ times}}(n).$$

3. Let f and g be arithmetic functions. Prove that $f * g = 0$ if and only if $f = 0$ or $g = 0$. It follows that the ring of arithmetic functions is an integral domain.

4. Let $\mathcal{A}$ be the ring of complex-valued arithmetic functions. An arithmetic function f is called a *unit* in $\mathcal{A}$ if there exists an arithmetic function g such that $f * g = \delta$. Prove that $f \in \mathcal{A}$ is a unit if and only if $f(1) \neq 0$.

5. For every positive integer N, let I_N be the set of all arithmetic functions $f(n)$ such that $f(n) = 0$ for all $n \leq N$. Prove that I_N is an ideal in the ring of arithmetic functions.

6. Let f and g be arithmetic functions. Prove that
$$L^n(f * g) = \sum_{k=0}^{n} \binom{n}{k} L^{n-k} f * L^k g.$$

7. Let $\mathcal{J}$ be the additive abelian semigroup consisting of all sequences $J = \{j_i\}_{i=1}^{\infty}$ of nonnegative integers such that $j_i = 0$ for all sufficiently large i. Addition of elements in $\mathcal{J}$ is defined coordinate-wise.

   Let $t_1, t_2, \ldots$ be an infinite sequence of variables. For every $J \in \mathcal{J}$ we define the monomial
$$t^J = \prod_{j_i \geq 1} t_i^{j_i}.$$
   If J is the sequence with $j_i = 0$ for all i, then $t^J = 1$. Let R be the set of all expressions of the form
$$\sum_{J \in \mathcal{J}} a_J t^J,$$
   where the coefficients a_j are complex numbers. We define the sum and product of elements of R by
$$\sum_{J \in \mathcal{J}} a_J t^J + \sum_{J \in \mathcal{J}} b_J t^J = \sum_{J \in \mathcal{J}} (a_J + b_J) t^J$$

and

$$\left(\sum_{J_1\in\mathcal{J}} a_{J_1}t^{J_1}\right)\left(\sum_{J_2\in\mathcal{J}} b_{J_2}t^{J_2}\right) = \sum_{J_1,J_2\in\mathcal{J}} a_{J_1}b_{J_2}t^{J_1+J_2}.$$

Prove that R is an integral domain, that is, a commutative ring with no zero divisors.

Remark. This ring is called the *ring of formal power series in infinitely many variables* $t_1, t_2 \ldots$ with coefficients in $\mathbf{C}$. It is denoted by $\mathbf{C}[[t_1, t_2, \ldots]]$.

8. Let $\mathbf{P} = \{p_1, p_2, p_3, \ldots\}$ be the sequence of primes in ascending order, that is, $p_1 = 2, p_2 = 3, p_3 = 5, \ldots$. By the fundamental theorem of arithmetic, to every positive integer n we can associate a sequence $J_n \in \mathcal{J}$ as follows: If

$$n = \prod_{i=1}^{\infty} p_i^{v_{p_i}(n)},$$

then

$$J_n = \{v_{p_i}(n)\}_{i=1}^{\infty}.$$

Prove that this is a bijection between $\mathbf{N}$ and $\mathcal{J}$.

9. Let $\mathcal{A}$ be the ring of complex-valued arithmetic functions. For every arithmetic function $f \in \mathcal{A}$ we define the formal power series

$$\Phi(f) = \sum_{n\in\mathbf{N}} f(n)t^{J_n} \in \mathbf{C}[[t_1, t_2, \ldots]],$$

where $J_n \in \mathcal{J}$ is the sequence constructed in Exercise 8. Prove that the map

$$\Phi : \mathcal{A} \to \mathbf{C}[[t_1, t_2, \ldots]]$$

is a ring isomorphism.

Remark. Since the ring of formal power series in infinitely many variables is a unique factorization domain, it follows that the ring of complex-valued arithmetic functions is also a unique factorization domain.

10. For arithmetic functions f and g, define the product $f \star g$ by

$$f \star g(n) = \sum_{k=1}^{n-1} f(k)g(n-k).$$

Is this product commutative? Is it associative? What is $f \star \delta$?

6.2 Mean Values of Arithmetic Functions

We define the *mean value* $F(x)$ of an arithmetic function $f(n)$ by

$$F(x) = \sum_{n \leq x} f(n),$$

where the sum is over all positive integers $n \leq x$. In particular, $F(x) = 0$ for $x < 1$. The function $F(x)$ is also called the *sum function* of f. We shall describe two simple but powerful tools for estimating sum functions in number theory. The first is integration and the second is partial summation.

The *integer part* of the real number x, denoted by $[x]$, is the unique integer n such that $n \leq x < n + 1$. The *fractional part* of x is the real number $\{x\} = x - [x] \in [0, 1)$. For example, $\left[-\frac{5}{3}\right] = -2$ and $\left\{-\frac{5}{3}\right\} = \frac{1}{3}$. Every real number x can be written uniquely in the form $x = [x] + \{x\}$.

A function $f(t)$ is *unimodal* on an interval I if there exists a number $t_0 \in I$ such that $f(t)$ is increasing for $t \leq t_0$ and decreasing for $t \geq t_0$. For example, the function $f(t) = \log^k t/t$ is unimodal on the interval $[1, \infty)$ with $t_0 = e^k$.

It is proved in real analysis that every function that is monotonic or unimodal on a closed interval $[a, b]$ is integrable.

Theorem 6.3 *Let a and b be integers with $a < b$, and let $f(t)$ be a function that is monotonic on the interval $[a, b]$. Then*

$$\min(f(a), f(b)) \leq \sum_{n=a}^{b} f(n) - \int_a^b f(t)dt \leq \max(f(a), f(b)). \tag{6.3}$$

Let x and y be real numbers with $y < [x]$, and let $f(t)$ be a nonnegative monotonic function on $[y, x]$. Then

$$\left| \sum_{y < n \leq x} f(n) - \int_y^x f(t)dt \right| \leq \max(f(y), f(x)). \tag{6.4}$$

If $f(t)$ is a nonnegative unimodal function on $[1, \infty)$, then

$$F(x) = \sum_{n \leq x} f(n) = \int_1^x f(t)dt + O(1). \tag{6.5}$$

Proof. If $f(t)$ is increasing on $[n, n+1]$, then

$$f(n) \leq \int_n^{n+1} f(t)dt \leq f(n+1).$$

If $f(t)$ is increasing on the interval $[a, b]$, then

$$f(a) + \int_a^b f(t)dt \leq \sum_{n=a}^{b} f(n) \leq f(b) + \int_a^b f(t)dt.$$

Similarly, if $f(t)$ is decreasing on the interval $[n, n+1]$, then

$$f(n+1) \leq \int_n^{n+1} f(t)dt \leq f(n).$$

If $f(t)$ is decreasing on the interval $[a, b]$, then

$$f(b) + \int_a^b f(t)dt \leq \sum_{n=a}^{b} f(n) \leq f(a) + \int_a^b f(t)dt.$$

This proves (6.3).

Let $f(t)$ be nonnegative and monotonic on the interval $[y, x]$. Let $a = [y] + 1$ and $b = [x]$. We have $y < a \leq b \leq x$. If $f(t)$ is increasing, then

$$\begin{aligned} \sum_{y<n\leq x} f(n) &= \sum_{a\leq n\leq b} f(n) \\ &\leq \int_a^b f(t)dt + f(b) \\ &\leq \int_y^x f(t)dt + f(x). \end{aligned}$$

Since

$$f(a) \geq \int_y^a f(t)dt$$

and

$$f(x) \geq \int_b^x f(t)dt,$$

it follows that

$$\begin{aligned} \sum_{y<n\leq x} f(n) &\geq \int_a^b f(t)dt + f(a) \\ &\geq \int_y^x f(t)dt - \int_b^x f(t)dt + f(a) - \int_y^a f(t)dt \\ &\geq \int_y^x f(t)dt - f(x). \end{aligned}$$

Therefore,

$$\left| \sum_{y<n\leq x} f(n) - \int_x^y f(t)dt \right| \leq f(x).$$

If $f(t)$ is decreasing, then

$$\sum_{y<n\leq x} f(n) = \sum_{a\leq n\leq b} f(n)$$

$$\begin{aligned} &\leq \int_a^b f(t)dt + f(a) \\ &\leq \int_y^x f(t)dt + f(y). \end{aligned}$$

Since

$$f(b) \geq \int_b^x f(t)dt$$

and

$$f(y) \geq \int_y^a f(t)dt,$$

it follows that

$$\begin{aligned} \sum_{y<n\leq x} f(n) &\geq \int_a^b f(t)dt + f(b) \\ &\geq \int_y^x f(t)dt + f(b) - \int_b^x f(t)dt - \int_y^a f(t)dt \\ &\geq \int_y^x f(t)dt - f(y) \end{aligned}$$

and

$$\left| \sum_{y<n\leq x} f(n) - \int_x^y f(t)dt \right| \leq f(y).$$

This proves (6.4).

If the function $f(t)$ is nonnegative and unimodal on $[1, \infty)$, then $f(t)$ is bounded and (6.5) follows from (6.4). □

Theorem 6.4 *For $x \geq 2$,*

$$\sum_{n\leq x} \log n = x \log x - x + O(\log x).$$

Proof. The function $f(t) = \log t$ is increasing on $[1, x]$. By Theorem 6.3,

$$\int_1^x \log t dt \leq \sum_{n\leq x} \log n \leq \int_1^x \log t dt + \log x,$$

and so

$$\sum_{n\leq x} \log n = x \log x - x + O(\log x).$$

This completes the proof. □

Theorem 6.5 *Let r be a nonnegative integer. For $x \geq 1$,*

$$\sum_{n \leq x} \frac{\log^r n}{n} = \frac{1}{r+1} \log^{r+1} x + O(1),$$

where the implied constant depends only on r.

Proof. The function $f(t) = \log^r t/t$ is nonnegative and unimodal on $[1, \infty)$ with maximum value $(r/e)^r$ at $t_0 = e^r$. By Theorem 6.3,

$$\sum_{n \leq x} \frac{\log^r n}{n} = \int_1^x \frac{\log^r t dt}{t} + O(1) = \frac{1}{r+1} \log^{r+1} x + O(1).$$

This completes the proof. □

Theorem 6.6 *Let k be a nonnegative integer. For $x \geq 1$,*

$$\sum_{n \leq x} \frac{\log^k (x/n)}{n} = \frac{1}{k+1} \log^{k+1} x + O(\log^k x),$$

where the implied constant depends only on k.

Proof. The idea is to expand $\log^k(x/n)$ by the binomial theorem and apply Theorem 6.5. We have

$$\begin{aligned}
\sum_{n \leq x} \frac{\log^k (x/n)}{n} &= \sum_{n \leq x} \frac{(\log x - \log n)^k}{n} \\
&= \sum_{n \leq x} \frac{1}{n} \sum_{r=0}^{k} \binom{k}{r} (-1)^r \log^{k-r} x \log^r n \\
&= \sum_{r=0}^{k} \binom{k}{r} (-1)^r \log^{k-r} x \sum_{n \leq x} \frac{\log^r n}{n} \\
&= \sum_{r=0}^{k} \binom{k}{r} (-1)^r \log^{k-r} x \left(\frac{1}{r+1} \log^{r+1} x + O(1) \right) \\
&= \sum_{r=0}^{k} \binom{k}{r} \frac{(-1)^r}{r+1} \log^{k+1} x + O\left(\sum_{r=0}^{k} \binom{k}{r} \log^{k-r} x \right) \\
&= \frac{1}{k+1} \log^{k+1} x + O(\log^k x),
\end{aligned}$$

since

$$\sum_{r=0}^{k} \frac{(-1)^r}{r+1} \binom{k}{r} = \frac{1}{k+1}$$

by Exercise 8. □

Theorem 6.7 *Let k be a positive integer. Then*

$$\sum_{n_1\cdots n_k\leq x}\frac{1}{n_1\cdots n_k}=\frac{1}{k!}\log^k x+O(\log^{k-1}x),$$

where $\sum_{n_1\cdots n_k\leq x}$ denotes the sum over all k-tuples of positive integers $(n_1,\ldots,n_k)$ such that $n_1\cdots n_k\leq x$.

Proof. By induction on k. For $k=1$, we set $r=0$ in Theorem 6.5 and obtain

$$\sum_{n_1\leq x}\frac{1}{n_1}=\log x+O(1).$$

Assume that the result holds for the positive integer k. Then

$$\begin{aligned}
&\sum_{n_1\cdots n_k n_{k+1}\leq x}\frac{1}{n_1\cdots n_k n_{k+1}}\\
&=\sum_{n_{k+1}\leq x}\frac{1}{n_{k+1}}\sum_{n_1\cdots n_k\leq x/n_{k+1}}\frac{1}{n_1\cdots n_k}\\
&=\sum_{n_{k+1}\leq x}\frac{1}{n_{k+1}}\left(\frac{1}{k!}\log^k(x/n_{k+1})+O(\log^{k-1}(x/n_{k+1}))\right)\\
&=\sum_{n_{k+1}\leq x}\frac{1}{k!n_{k+1}}(\log x-\log n_{k+1})^k\\
&\quad+O\left(\log^{k-1}x\sum_{n_{k+1}\leq x}\frac{1}{n_{k+1}}\right)\\
&=\sum_{n\leq x}\frac{1}{k!n}(\log x-\log n)^k+O\left(\log^k x\right).
\end{aligned}$$

We use the binomial theorem and Theorem 6.5 to compute the main term.

$$\begin{aligned}
\sum_{n\leq x}\frac{1}{k!n}(\log x-\log n)^k &= \sum_{n\leq x}\frac{1}{k!n}\sum_{r=0}^{k}(-1)^r\binom{k}{r}\log^{k-r}x\log^r n\\
&=\sum_{r=0}^{k}\frac{(-1)^r}{k!}\binom{k}{r}\log^{k-r}x\sum_{n\leq x}\frac{\log^r n}{n}\\
&=\sum_{r=0}^{k}\frac{(-1)^r}{k!}\binom{k}{r}\log^{k-r}x\left(\frac{1}{r+1}\log^{r+1}x+O(1)\right)
\end{aligned}$$

$$= \frac{1}{k!} \log^{k+1} x \sum_{r=0}^{k} \frac{(-1)^r}{r+1} \binom{k}{r} + O\left(\log^k x\right)$$
$$= \frac{1}{(k+1)!} \log^{k+1} x + O\left(\log^k x\right),$$

by Exercise 8. □

Theorem 6.8 (Partial summation) *Let $f(n)$ and $g(n)$ be arithmetic functions. Consider the sum function*

$$F(x) = \sum_{n \leq x} f(n).$$

Let a and b be nonnegative integers with $a < b$. Then

$$\sum_{n=a+1}^{b} f(n)g(n) = F(b)g(b) - F(a)g(a+1) - \sum_{n=a+1}^{b-1} F(n)(g(n+1) - g(n)). \tag{6.6}$$

Let x and y be nonnegative real numbers with $[y] < [x]$, and let $g(t)$ be a function with a continuous derivative on the interval $[y, x]$. Then

$$\sum_{y < n \leq x} f(n)g(n) = F(x)g(x) - F(y)g(y) - \int_y^x F(t)g'(t)dt. \tag{6.7}$$

In particular, if $x \geq 2$ and $g(t)$ is continuously differentiable on $[1, x]$, then

$$\sum_{n \leq x} f(n)g(n) = F(x)g(x) - \int_1^x F(t)g'(t)dt. \tag{6.8}$$

Proof. Identity (6.6) is a straightforward calculation:

$$\begin{aligned}
&\sum_{n=a+1}^{b} f(n)g(n) \\
&= \sum_{n=a+1}^{b} (F(n) - F(n-1))g(n) \\
&= \sum_{n=a+1}^{b} F(n)g(n) - \sum_{n=a}^{b-1} F(n)g(n+1) \\
&= F(b)g(b) - F(a)g(a+1) - \sum_{n=a+1}^{b-1} F(n)(g(n+1) - g(n)).
\end{aligned}$$

If the function $g(t)$ is continuously differentiable on $[y, x]$, then

$$g(n+1) - g(n) = \int_n^{n+1} g'(t)dt.$$

Since $F(t) = F(n)$ for $n \leq t < n+1$, it follows that

$$F(n)\left(g(n+1) - g(n)\right) = \int_n^{n+1} F(t)g'(t)dt.$$

Let $a = [y]$ and $b = [x]$. Since $a \leq y < a+1 \leq b \leq x < b+1$, we have

$$\begin{aligned}
&\sum_{y<n\leq x} f(n)g(n) \\
&= \sum_{n=a+1}^{b} f(n)g(n) \\
&= F(b)g(b) - F(a)g(a+1) - \sum_{n=a+1}^{b-1} F(n)(g(n+1) - g(n)) \\
&= F(x)g(b) - F(y)g(a+1) - \sum_{n=a+1}^{b-1} \int_n^{n+1} F(t)g'(t)dt \\
&= F(x)g(x) - F(y)g(y) - F(x)(g(x) - g(b)) - F(y)(g(a+1) - g(y)) \\
&\quad - \int_{a+1}^{b} F(t)g'(t)dt \\
&= F(x)g(x) - F(y)g(y) - \int_y^x F(t)g'(t)dt.
\end{aligned}$$

This proves (6.7).

If $x \geq 2$ and $g(t)$ is continuously differentiable on $[1, x]$, then

$$\begin{aligned}
\sum_{n\leq x} f(n)g(n) &= f(1)g(1) + \sum_{1<n\leq x} f(n)g(n) \\
&= f(1)g(1) + F(x)g(x) - F(1)g(1) - \int_1^x F(t)g'(t)dt \\
&= F(x)g(x) - \int_1^x F(t)g'(t)dt.
\end{aligned}$$

This proves (6.8). □

Letting $r = 0$ in Theorem 6.5, we obtain $\sum_{n\leq x} 1/n = \log x + O(1)$. Using partial summation, we can obtain a more precise result.

Theorem 6.9 *For* $x \geq 1$,

$$\sum_{n \leq x} \frac{1}{n} = \log x + \gamma + r(x),$$

where

$$0 < \gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2}\,dt < 1$$

and

$$|r(x)| < \frac{1}{x}.$$

The number $\gamma = 0.577\ldots$ is called *Euler's constant.* A famous unsolved problem in number theory is to determine whether γ is rational or irrational.

Proof. Since $0 \leq \{t\} < 1$ for all t, we have

$$0 < \int_1^\infty \frac{\{t\}}{t^2}\,dt < \int_1^\infty \frac{1}{t^2}\,dt = 1,$$

and so $\gamma \in (0, 1)$.

We apply partial summation to the functions $f(n) = 1$ and $g(t) = 1/t$. Then $F(t) = \sum_{n \leq t} 1 = [t]$ and

$$\begin{aligned}
\sum_{n \leq x} \frac{1}{n} &= \sum_{n \leq x} f(n)g(n) \\
&= \frac{[x]}{x} + \int_1^x \frac{[t]}{t^2}\,dt \\
&= 1 - \frac{\{x\}}{x} + \int_1^x \frac{1}{t}\,dt - \int_1^x \frac{\{t\}}{t^2}\,dt \\
&= \log x + \left(1 - \int_1^\infty \frac{\{t\}}{t^2}\,dt\right) + \int_x^\infty \frac{\{t\}}{t^2}\,dt - \frac{\{x\}}{x} \\
&= \log x + \gamma + r(x),
\end{aligned}$$

where

$$r(x) = \int_x^\infty \frac{\{t\}}{t^2}\,dt - \frac{\{x\}}{x}.$$

Moreover, $|r(x)| < 1/x$ since $0 \leq \{x\}/x < 1$ and

$$0 < \int_x^\infty \frac{\{t\}}{t^2}\,dt < \int_x^\infty \frac{1}{t^2}\,dt = \frac{1}{x}.$$

□

Theorem 6.10 *Let $A = \{a_i\}_{i=1}^{\infty}$ be an infinite set of positive integers with $a_1 < a_2 < a_3 < \cdots$. If*

$$A(x) = \sum_{a_i \leq x} 1 = O\left(\frac{x}{\log^2 x}\right)$$

for $x \geq 2$, then the series

$$\sum_{i=1}^{\infty} \frac{1}{a_i}$$

converges.

Proof. Let $\chi_A(n)$ be the characteristic function of A, that is,

$$\chi_A(n) = \begin{cases} 1 & \text{if } n \in A, \\ 0 & \text{if } n \notin A. \end{cases}$$

There exists a number c such that

$$A(x) = \sum_{n \leq x} \chi_A(n) \leq \frac{cx}{\log^2 x}$$

for all $x \geq 2$, and $A(x) \leq 1$ for $1 \leq x < 2$. Applying partial summation, we obtain

$$\begin{aligned} \sum_{a_i \leq x} \frac{1}{a_i} &= \sum_{n \leq x} \frac{\chi_A(n)}{n} \\ &= \frac{A(x)}{x} + \int_1^x \frac{A(t)dt}{t^2} \\ &\leq \frac{c}{\log^2 x} + \frac{1}{2} + c\int_2^x \frac{dt}{t \log^2 t} \\ &= \frac{c}{\log^2 x} + \frac{1}{2} + c\int_{\log 2}^{\log x} \frac{du}{u^2} \\ &< \infty. \end{aligned}$$

This completes the proof. □

Theorem 6.11 *For $x \geq 2$,*

$$\sum_{n \leq x} \log^2 n = x \log^2 x - 2x \log x + 2x + O\left(\log^2 x\right).$$

Proof. We use partial summation with $f(n) = 1$ and $g(t) = \log^2 t$. Then $F(t) = [t]$ and $g'(t) = 2\log t/t$. Then

$$\begin{aligned}
\sum_{n\le x} \log^2 n &= [x]\log^2 x - 2\int_1^x \frac{[t]\log t}{t}dt \\
&= (x - \{x\})\log^2 x - 2\int_1^x \frac{(t-\{t\})\log t}{t}dt \\
&= x\log^2 x + O(\log^2 x) - 2\int_1^x \log t dt + 2\int_1^x \frac{\{t\}\log t}{t}dt \\
&= x\log^2 x - 2x\log x + 2x + O(\log^2 x).
\end{aligned}$$

This completes the proof. □

Theorem 6.12 *For* $x \ge 2$,

$$\sum_{n\le x} \log^2 \frac{x}{n} = 2x + O\left(\log^2 x\right).$$

Proof. From Theorem 6.4 and Theorem 6.11, we obtain

$$\begin{aligned}
\sum_{n\le x} \log^2 \frac{x}{n} &= \sum_{n\le x} (\log x - \log n)^2 \\
&= \sum_{n\le x} (\log^2 x - 2\log x\log n + \log^2 n) \\
&= [x]\log^2 x - 2\log x \sum_{n\le x}\log n + \sum_{n\le x}\log^2 n \\
&= x\log^2 x - 2\log x(x\log x - x) + x\log^2 x - 2x\log x + 2x + O\left(\log^2 x\right) \\
&= 2x + O\left(\log^2 x\right).
\end{aligned}$$

This completes the proof. □

Exercises

1. Prove that
$$e\left(\frac{n}{e}\right)^n < n! < en\left(\frac{n}{e}\right)^n.$$
Hint: Use partial summation to estimate $\log n!$.

2. Let $f(n)$ be an arithmetic function such that
$$F(x) = \sum_{n\le x} f(n) = O(x).$$

Prove that

$$\sum_{n\leq x} \frac{f(n)}{n} = O(\log x).$$

3. Prove that

$$\sum_{n\leq x} \frac{1}{n^{1/2}} = 2x^{1/2} - \left(1 + \int_1^\infty \frac{\{t\}}{2t^{3/2}}dt\right) + O\left(x^{-1/2}\right).$$

4. For $0 < a < 1$, let

$$\gamma(a) = \frac{a}{1-a} + a\int_1^\infty \frac{\{t\}}{t^{a+1}}dt.$$

Prove that

$$\sum_{n\leq x} \frac{1}{n^a} = \frac{x^{1-a}}{1-a} - \gamma(a) + O\left(x^{-a}\right).$$

5. Prove that

$$\sum_{n\leq x} \log^k n = x\log^k x + O(x\log^{k-1} x)$$

for all positive integers k.

6. Prove that

$$\sum_{n\leq x} \log \frac{x}{n} = x + O(\log x).$$

7. Prove that

$$\sum_{n\leq x} \log^k \frac{x}{n} = k!x + O(\log^k x)$$

for all positive integers k.

8. Prove that for every nonnegative integer k,

$$\sum_{r=0}^{k} \frac{(-1)^r}{r+1}\binom{k}{r} = \frac{1}{k+1}.$$

9. Prove that for every positive integer j,

$$\sum_{n=1}^{r} n^j = \frac{r^{j+1}}{j+1} + O(r^j).$$

10. Let a, b and k be positive integers, with $a < b$ and $k \geq 2$. Prove that

$$\sum_{n=a}^{b} \frac{1}{n^2} = \left(\frac{1}{b} - \frac{1}{a}\right) + O\left(\frac{1}{a^2}\right).$$

Prove that

$$\sum_{n=a}^{b} \frac{1}{n^k} = \frac{1}{k-1}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right) + O\left(\frac{1}{a^k}\right).$$

11. Prove that

$$\sum_{n \leq x} \frac{1}{1 + n\log n} = O(\log\log x).$$

6.3 The Möbius Function

The *Möbius function* $\mu(n)$ is defined as follows:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes,} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases}$$

We have

$$\begin{array}{rclcrcl} \mu(1) & = & 1, & \quad & \mu(6) & = & 1, \\ \mu(2) & = & -1, & & \mu(7) & = & -1, \\ \mu(3) & = & -1, & & \mu(8) & = & 0, \\ \mu(4) & = & 0, & & \mu(9) & = & 0, \\ \mu(5) & = & -1, & & \mu(10) & = & 1. \end{array}$$

An integer is called *square-free* if it is not divisible by the square of a prime. Thus, $\mu(n) \neq 0$ if and only if n is square-free.

Recall that an arithmetic function $f(n)$ is *multiplicative* if $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$.

Theorem 6.13 *The Möbius function $\mu(n)$ is multiplicative, and*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases} \tag{6.9}$$

Proof. Multiplicativity follows immediately from the definition of the Möbius function, since if m and n are relatively prime square-free integers with k and ℓ prime factors, respectively, then mn is square-free with $k + \ell$ factors, and

$$\mu(m)\mu(n) = (-1)^k(-1)^\ell = (-1)^{k+\ell} = \mu(mn).$$

Next we prove the convolution formula (6.9). If $n = 1$, then

$$\sum_{d|n} \mu(d) = \mu(1) = 1.$$

For $n \geq 2$, let

$$n = p_1^{r_1} \cdots p_k^{r_k}$$

be the standard factorization of the integer n. Then $r \geq 1$. Recall that the *radical* of n is the largest square-free divisor of n, that is,

$$\mathrm{rad}(n) = p_1 \cdots p_r$$

is the product of the distinct primes dividing n. Let $m = \mathrm{rad}(n)$. If d divides n and $\mu(d) \neq 0$, then d is square-free, and so d divides m. Since m is the product of k primes, it follows that there are exactly $\binom{k}{i}$ divisors of m that can be written as the product of i distinct primes, that is, the number of divisors d of m such that $\omega(d) = i$ is $\binom{k}{i}$. Therefore,

$$\begin{aligned}
\sum_{d|n} \mu(d) &= \sum_{d|m} \mu(d) \\
&= \sum_{i=0}^{k} \sum_{\substack{d|m \\ \omega(d)=i}} \mu(d) \\
&= \sum_{i=0}^{k} \sum_{\substack{d|m \\ \omega(d)=i}} (-1)^i \\
&= \sum_{i=0}^{k} \binom{k}{i} (-1)^i \\
&= (1-1)^k \\
&= 0.
\end{aligned}$$

This completes the proof.

We defined the arithmetic function $1(n)$ by $1(n) = 1$ for all n. Using the Dirichlet convolution, we can restate Theorem 6.13 as follows:

$$\mu * 1 = \delta,$$

and so the Möbius function μ is a unit with inverse 1.

Theorem 6.14 (Möbius inversion) *If f is any arithmetic function, and g is the arithmetic function defined by*

$$g(n) = \sum_{d|n} f(d),$$

then

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d).$$

Similarly, if g is any arithmetic function, and f is the arithmetic function defined by

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d),$$

then

$$g(n) = \sum_{d|n} f(d).$$

Proof. We use Theorem 6.13 and the commutativity and associativity of Dirichlet convolution. The definition

$$g(n) = \sum_{d|n} f(d)$$

is equivalent to

$$g = f * 1.$$

Then

$$g * \mu = (f * 1) * \mu = f * (1 * \mu) = f * \delta = f.$$

Similarly, if

$$f = g * \mu,$$

then

$$f * 1 = (g * \mu) * 1 = g * (\mu * 1) = g * \delta = g.$$

This completes the proof. □

The following result gives a useful identity for sum functions of arithmetic functions. The proof can be described geometrically as a sum over the lattice points (m, d) under the hyperbola $v = x/u$ in the positive quadrant of the uv-plane.

Theorem 6.15 *Let $f(n)$ be an arithmetic function and*

$$F(x) = \sum_{n \leq x} f(n).$$

Then

$$\sum_{m \leq x} F\left(\frac{x}{m}\right) = \sum_{d \leq x} f(d) \left[\frac{x}{d}\right] = \sum_{n \leq x} \sum_{d|n} f(d).$$

Proof. We have

$$\begin{aligned}\sum_{m\le x} F\left(\frac{x}{m}\right) &= \sum_{m\le x}\sum_{d\le x/m} f(d) = \sum_{dm\le x} f(d)\\ &= \sum_{d\le x} f(d) \sum_{m\le x/d} 1 = \sum_{d\le x} f(d)\left[\frac{x}{d}\right].\\ &= \sum_{n\le x}\sum_{d\mid n} f(d).\end{aligned}$$

Also,

$$\sum_{m\le x} F\left(\frac{x}{m}\right) = \sum_{dm\le x} f(d) = \sum_{n\le x}\sum_{d\mid n} f(d).$$

This completes the proof. □

Theorem 6.16

$$\sum_{n\le x}\frac{\mu(n)}{n} = O(1).$$

Proof. Applying Theorem 6.15 with $f(n) = \mu(n)$ and

$$M(x) = \sum_{n\le x}\mu(n),$$

we obtain

$$\sum_{m\le x} M\left(\frac{x}{m}\right) = \sum_{d\le x}\mu(d)\left[\frac{x}{d}\right] = \sum_{n\le x}\sum_{d\mid n}\mu(d) = 1,$$

by Theorem 6.13. Since

$$\sum_{d\le x}\mu(d)\left[\frac{x}{d}\right] = x\sum_{d\le x}\frac{\mu(d)}{d} - \sum_{d\le x}\mu(d)\left\{\frac{x}{d}\right\} = x\sum_{d\le x}\frac{\mu(d)}{d} + O(x),$$

it follows that

$$x\sum_{d\le x}\frac{\mu(d)}{d} + O(x) = 1.$$

Therefore,

$$x\sum_{d\le x}\frac{\mu(d)}{d} = O(x),$$

and so

$$\sum_{d\le x}\frac{\mu(d)}{d} = O(1).$$

This completes the proof. □

Theorem 6.17

$$\sum_{n \leq x} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2} + O\left(\frac{1}{x}\right).$$

Proof. The *Riemann zeta function*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

converges absolutely for $s > 1$. Similarly, the function

$$G(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

converges absolutely for $s > 1$. Therefore,

$$\begin{aligned}
\zeta(s)G(s) &= \sum_{k=1}^{\infty} \frac{1}{k^s} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^s} \\
&= \sum_{k=1}^{\infty} \sum_{d=1}^{\infty} \frac{\mu(d)}{(kd)^s} \\
&= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{d \mid n} \mu(d) \\
&= 1,
\end{aligned}$$

by Theorem 6.13, and so

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

for $s > 1$. Since

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

it follows that

$$\frac{1}{\zeta(2)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2},$$

and so

$$\left| \sum_{n \leq x} \frac{\mu(n)}{n^2} - \frac{6}{\pi^2} \right| = \left| \sum_{n > x} \frac{\mu(n)}{n^2} \right| < \sum_{n > x} \frac{1}{n^2} \ll \frac{1}{x}.$$

This completes the proof. □

Exercises

1. Compute $\mu(n)$ for $11 \leq n \leq 30$.

2. Let $f(n)$ be an arithmetic function, and define $g(n) = \sum_{d|n} f(d)$. Use Möbius inversion to write $f(30)$ as a sum and difference of values of the arithmetic function g.

3. Let $d(n)$ be the divisor function. Prove that
$$\sum_{k|n} d(k)\mu\left(\frac{n}{k}\right) = 1$$
for every positive integer n.

   Hint: Problem 1 in Section 6.1.

4. Let $\sigma(n)$ denote the sum of the positive divisors of n, that is,
$$\sigma(n) = \sum_{k|n} k.$$
Prove that
$$\sum_{k|n} \sigma(k)\mu\left(\frac{n}{k}\right) = n$$
for every positive integer n.

5. Let $f(x)$ be a function on the set of real numbers $x \geq 1$. Define the function $g(x)$ by
$$g(x) = \sum_{n \leq x} f\left(\frac{x}{n}\right).$$
Prove that
$$f(x) = \sum_{n \leq x} \mu(n) g\left(\frac{x}{n}\right).$$

6. Let $g(x)$ be a function on the set of real numbers $x \geq 1$. Define the function $f(x)$ by
$$f(x) = \sum_{n \leq x} \mu(n) g\left(\frac{x}{n}\right).$$
Prove that
$$g(x) = \sum_{n \leq x} f\left(\frac{x}{n}\right).$$

7. Let $\alpha > 0$. Let $f(x)$ be a function on the set of real numbers $x \geq 1$. Define the function $g(x)$ by
$$g(x) = \sum_{n \leq x^{1/\alpha}} \frac{1}{n^\alpha} f\left(\frac{x}{n^\alpha}\right).$$

Prove that

$$f(x) = \sum_{n \leq x^{1/\alpha}} \frac{\mu(n)}{n^\alpha} g\left(\frac{x}{n^\alpha}\right).$$

8. Let $\alpha > 0$. Let $g(x)$ be a function on the set of real numbers $x \geq 1$. Define the function $f(x)$ by

$$f(x) = \sum_{n^{1/\alpha} \leq x} \frac{\mu(n)}{n^\alpha} g\left(\frac{x}{n^\alpha}\right).$$

Prove that

$$g(x) = \sum_{n \leq x^{1/\alpha}} \frac{1}{n^\alpha} f\left(\frac{x}{n^\alpha}\right).$$

9. Prove that every positive integer n can be written uniquely in the form $n = k^2\ell$, where k and ℓ are positive integers and ℓ is square-free. Prove that

$$\mu^2(n) = \sum_{d^2 | n} \mu(d).$$

10. Prove that the density of the square-free integers is $6/\pi^2$. Equivalently, let $Q(x)$ denote the number of square-free integers not exceeding x. Prove that

$$\lim_{x \to \infty} \frac{Q(x)}{x} = \frac{6}{\pi^2}.$$

Hint: n is square-free if and only if $\mu^2(n) = 1$, and

$$Q(x) = \sum_{n \leq x} \mu^2(n) = \sum_{d^2 \leq x} \mu(d) \left[\frac{x}{d^2}\right] = \frac{6x}{\pi^2} + O(\sqrt{x}).$$

11. Define the von Mangoldt function

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ is a prime power,} \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$L(n) = \log n.$$

Prove that

$$L = 1 * \Lambda$$

and

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d.$$

6.4 Multiplicative Functions

In this section we prove some general properties about multiplicative arithmetic functions.

Theorem 6.18 *If f is a multiplicative function, then*

$$f([m,n])f((m,n)) = f(m)f(n)$$

for all positive integers m and n.

Proof. Let $p_1, \ldots, p_r$ be the prime numbers that divide m or n. Then

$$n = \prod_{i=1}^{r} p_i^{k_i}$$

and

$$m = \prod_{i=1}^{r} p_i^{\ell_i},$$

where $k_1, \ldots, k_r, \ell_1, \ldots, \ell_r$ are nonnegative integers. Then

$$[m,n] = \prod_{i=1}^{r} p_i^{\max(k_i,\ell_i)}$$

and

$$(m,n) = \prod_{i=1}^{r} p_i^{\min(k_i,\ell_i)}.$$

Since

$$\{\max(k_i,\ell_i), \min(k_i,\ell_i)\} = \{k_i, \ell_i\}$$

and since f is multiplicative, it follows that

$$\begin{aligned} f([m,n])f((m,n)) &= \prod_{i=1}^{r} f\left(p_i^{\max(k_i,\ell_i)}\right) \prod_{i=1}^{r} f\left(p_i^{\min(k_i,\ell_i)}\right) \\ &= \prod_{i=1}^{r} f(p_i^{k_i}) \prod_{i=1}^{r} f(p_i^{\ell_i}) \\ &= f(m)f(n). \end{aligned}$$

This completes the proof. □

Theorem 6.19 *Let f be a multiplicative function with $f(1) = 1$. Then*

$$\sum_{d|n} \mu(d)f(d) = \prod_{p|n} (1 - f(p)).$$

Proof. The identity holds for $n = 1$. For $n \geq 2$, let $m = \mathrm{rad}(n)$ be the product of the distinct primes dividing n. Since $\mu(d) = 0$ if d is not square-free, it follows that

$$\sum_{d|n} \mu(d) f(d) = \sum_{d|m} \mu(d) f(d) = \prod_{p|m}(1 - f(p)) = \prod_{p|n}(1 - f(p)).$$

This completes the proof. □

The sequence of prime powers is the sequence

$$2, 3, 4, 5, 7, 8, 9, 11, 13, 16, 17, 19, 23, 25, 27, \ldots .$$

The smallest power that is not a prime power is 36.

Theorem 6.20 *Let $f(n)$ be a multiplicative function. If*

$$\lim_{p^k \to \infty} f(p^k) = 0$$

as p^k runs through the sequence of all prime powers, then

$$\lim_{n \to \infty} f(n) = 0.$$

Proof. Since $\lim_{p^k \to \infty} f(p^k) = 0$, it follows that there exist only finitely many prime powers p^k such that $|f(p^k)| \geq 1$, and so we can define

$$A = \prod_{|f(p^k)| \geq 1} |f(p^k)|.$$

Then $A \geq 1$.

Let $0 < \varepsilon < 1$. There exist only finitely many prime powers p^k such that $|f(p^k)| \geq \varepsilon / A$, and so there are only finitely many integers n such that

$$|f(p^k)| \geq \frac{\varepsilon}{A}$$

for every prime power p^k that exactly divides n. Therefore, if n is sufficiently large, then n is divisible by at least one prime power p^k such that $|f(p^k)| < \varepsilon / A$, and so n can be written in the form

$$n = \prod_{i=1}^{r} p_i^{k_i} \prod_{i=r+1}^{r+s} p_i^{k_i} \prod_{i=r+s+1}^{r+s+t} p_i^{k_i},$$

where $p_1, \ldots, p_{r+s+t}$ are distinct prime numbers such that

$$|f(p_i^{k_i}| \geq 1 \qquad \text{for } i = 1, \ldots, r,$$

$$\frac{\varepsilon}{A} \leq |f(p_i^{k_i}| < 1 \qquad \text{for } i = r+1, \ldots, r+s,$$

$$|f(p_i^{k_i}| < \frac{\varepsilon}{A} \qquad \text{for } i = r+s+1, \ldots, r+s+t,$$

and

$$t \geq 1.$$

Since f is multiplicative,

$$|f(n)| = \prod_{i=1}^{r} |f(p_i^{k_i})| \prod_{i=r+1}^{r+s} |f(p_i^{k_i})| \prod_{i=r+s+1}^{r+s+t} |f(p_i^{k_i})| < A(\varepsilon/A)^t \leq \varepsilon.$$

This completes the proof. □

Exercises

1. Let f be a multiplicative function. Prove that if $f(1) = 0$, then f is identically equal to 0, that is, $f(n) = 0$ for all n. Prove that if f is not identically equal to 0, then $f(1) = 1$.

2. Prove that a multiplicative function is completely determined by its values on prime powers p^k.

3. Prove that if f and g are multiplicative functions, then $f * g$ is also multiplicative.

4. Define the arithmetic functions $\omega(n)$ and $\Omega(n)$ as follows: If

   $$n = p_1^{k_1} \cdots p_r^{k_r}$$

   is the standard factorization of the positive integer n, then

   $$\omega(n) = r$$

   is the number of distinct prime divisors of n, and

   $$\Omega(n) = k_1 + \cdots + k_r$$

   is the total number of prime factors of n. Prove that n is square-free if and only if $\omega(n) = \Omega(n)$. Prove that the arithmetic function $(-1)^{\omega(n)}$ is multiplicative.

5. An arithmetic function f is called *completely multiplicative* if $f(mn) = f(m)f(n)$ for all positive integers m and n. Prove that Liouville's function

   $$\lambda(n) = (-1)^{\Omega(n)}$$

is completely multiplicative. Prove that

$$\sum_{d|n} \lambda(d) = \begin{cases} 1 & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

6. Prove that for every $\delta > 0$,

$$\lim_{n\to\infty} \frac{\varphi(n)}{n^{1-\delta}} = \infty.$$

Hint: Apply Theorem 6.20 to the multiplicative function $f(n) = n^{1-\delta}/\varphi(n)$. Observe that

$$0 < \frac{p^{k(1-\delta)}}{p^k(1-p^{-1})} \leq \frac{2}{p^{k\delta}}.$$

7. Prove that

$$\prod_{p|n}\left(1-\frac{1}{p^2}\right) \geq \prod_{k=2}^{n}\left(1-\frac{1}{k^2}\right) > \frac{1}{2}.$$

Hint: Consider the identity

$$\prod_{k=2}^{n}\left(1-\frac{1}{k^2}\right) = \prod_{k=2}^{n}\left(\frac{k-1}{k}\right)\prod_{k=2}^{n}\left(\frac{k+1}{k}\right).$$

8. Prove that

$$\frac{1}{2} < \frac{\varphi(n)\sigma(n)}{n^2} < 1.$$

Hint: Observe that for every prime power p^k,

$$\frac{\varphi(p^k)\sigma(p^k)}{p^{2k}} = 1 - \frac{1}{p^{k+1}} \geq 1 - \frac{1}{p^2}.$$

9. Prove that

$$n < \sigma(n) \ll n^{1+\delta}$$

for every $\delta > 0$.

Hint: Apply Exercise 6 and Exercise 8.

6.5 The mean value of the Euler Phi Function

The Euler phi function is

$$\varphi(n) = n\prod_{p|n}\left(1-\frac{1}{p}\right) = n\sum_{d|n}\frac{\mu(d)}{d} = \sum_{d'd=n} d'\mu(d). \tag{6.10}$$

We shall find an asymptotic formula for the mean value of the Euler phi function.

Theorem 6.21 *For* $x \geq 1$,

$$\Phi(x) = \sum_{n \leq x} \varphi(n) = \frac{3x^2}{\pi^2} + O(x \log x).$$

Proof. We have

$$\begin{aligned}
\Phi(x) &= \sum_{n \leq x} \varphi(n) \\
&= \sum_{n \leq x} \sum_{d'd=n} d' \mu(d) \\
&= \sum_{d \leq x} \mu(d) \sum_{d' \leq x/d} d' \\
&= \frac{1}{2} \sum_{d \leq x} \mu(d) \left[\frac{x}{d}\right] \left(\left[\frac{x}{d}\right] + 1\right) \\
&= \frac{1}{2} \sum_{d \leq x} \mu(d) \left(\left(\frac{x}{d}\right)^2 + O\left(\frac{x}{d}\right)\right) \\
&= \frac{x^2}{2} \sum_{d \leq x} \frac{\mu(d)}{d^2} + O\left(x \sum_{d \leq x} \frac{1}{d}\right) \\
&= \frac{x^2}{2} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \frac{x^2}{2} \sum_{d > x} \frac{\mu(d)}{d^2} + O(x \log x) \\
&= \frac{3x^2}{\pi^2} + O(x \log x).
\end{aligned}$$

This completes the proof. □

Theorem 6.22 *The probability that two positive integers are relatively prime is* $6/\pi^2$.

Proof. Let $N \geq 1$. The number of ordered pairs of positive integers (m, n) such that $1 \leq m \leq n \leq N$ is $N + \binom{N}{2} = N(N+1)/2$. The number of positive integers $m \leq n$ that are relatively prime is $\varphi(n)$, and so the number of pairs of positive integers (m, n) such that $1 \leq m \leq n \leq N$ and m and n are relatively prime is

$$\sum_{n \leq N} \varphi(n) = \frac{3N^2}{\pi^2} + O(N \log N).$$

Therefore, the frequency of relatively prime pairs of positive integers not exceeding N is

$$\frac{\frac{3N^2}{\pi^2} + O(N \log N)}{N(N+1)/2} = \frac{6}{\pi^2} + O\left(\frac{\log N}{N}\right) \longrightarrow \frac{6}{\pi^2}$$

as $N \to \infty$. This completes the proof.

Exercises

1. Use Möbius inversion to prove identity (6.10):

$$\varphi(n) = n \sum_{d|n} \frac{\mu(d)}{d}.$$

2. Prove that

$$\limsup_{n\to\infty} \frac{\varphi(n)}{n} = 1.$$

   Hint: Consider $\varphi(n)$ for $n = p$ prime.

6.6 Notes

Everything in this chapter is classical number theory. For other elementary results on arithmetic functions, see Hardy and Wright [60].

There is a vast literature on the distribution of values of arithmetic functions. For a comprehensive survey of this field, see Elliott, *Probabilistic Number Theory I, II* [28, 29].

7
Divisor Functions

7.1 Divisors and Factorizations

The *divisor function* $d(n)$ counts the number of positive divisors of n. Thus,

$$\begin{array}{rclcrcl} d(1) & = & 1, & \quad & d(6) & = & 4, \\ d(2) & = & 2, & & d(7) & = & 2, \\ d(3) & = & 2, & & d(8) & = & 4, \\ d(4) & = & 3, & & d(9) & = & 3, \\ d(5) & = & 2, & & d(10) & = & 4. \end{array}$$

We can write down an explicit formula for $d(n)$ in terms of the prime powers that exactly divide n. Let

$$n = \prod_{p|n} p^{v_p(n)}.$$

Every divisor d of n is of the form

$$d = \prod_{p|n} p^{a_p},$$

where a_p is an integer such that

$$0 \leq a_p \leq v_p(n).$$

Since each exponent a_p can be chosen in $v_p(n) + 1$ ways, it follows that

$$d(n) = \prod_{p|n} (v_p(n) + 1).$$

Theorem 7.1 *The divisor function $d(n)$ is multiplicative.*

Proof. Let m and n be relatively prime integers,

$$m = \prod_{p|m} p^{v_p(m)}$$

and

$$n = \prod_{q|n} q^{v_q(n)}.$$

Since $(m,n) = 1$, the set of primes that divide m and the set of primes that divide n are disjoint. Therefore,

$$mn = \prod_{p|m} p^{v_p(m)} \prod_{q|n} q^{v_q(n)}$$

is the standard factorization of mn, and

$$d(mn) = \prod_{p|m}(v_p(m)+1)\prod_{q|n}(v_q(n)+1) = d(m)d(n).$$

This completes the proof. □

Theorem 7.2 *For every $\varepsilon > 0$,*

$$d(n) \ll_\varepsilon n^\varepsilon.$$

Proof. Let $\varepsilon > 0$. The function $f(n) = d(n)/n^\varepsilon$ is multiplicative. Therefore, by Theorem 6.20, it suffices to prove that

$$\lim_{p^k\to\infty} f(p^k) = 0$$

for every prime p. We observe that

$$\frac{k+1}{2^{k\varepsilon/2}}$$

is bounded for $k \geq 1$, and so

$$\begin{aligned} f(p^k) &= \frac{d(p^k)}{p^{k\varepsilon}} \\ &= \frac{k+1}{p^{k\varepsilon}} \\ &= \left(\frac{k+1}{p^{k\varepsilon/2}}\right)\left(\frac{1}{p^{k\varepsilon/2}}\right) \\ &\leq \left(\frac{k+1}{2^{k\varepsilon/2}}\right)\left(\frac{1}{p^{k\varepsilon/2}}\right) \\ &\ll \frac{1}{p^{k\varepsilon/2}}. \end{aligned}$$

This completes the proof. □

Theorem 7.3 *For $x \geq 1$,*

$$D(x) = \sum_{n \leq x} d(n) = x \log x + (2\gamma - 1)x + O(\sqrt{x}).$$

The problem of estimating the sum function $D(x)$ is called *Dirichlet's divisor problem.*

Proof. We can interpret the divisor function $d(n)$ and the sum function $D(x)$ geometrically. A *lattice point* in the plane is a point whose coordinates are integers. A *positive lattice point* in the plane is a point whose coordinates are positive integers. In the uv-plane,

$$d(n) = \sum_{d|n} 1 = \sum_{n=uv} 1$$

counts the number of lattice points (u, v) on the rectangular hyperbola $uv = n$ that lie in the quadrant $u > 0, v > 0$. The sum function $D(x)$ counts the number of lattice points in this quadrant that lie on or under the hyperbola $uv = x$, that is, the number of positive lattice points (u, v) such that $1 \leq u \leq x$ and $1 \leq v \leq x/u$. These lattice points can be divided into three pairwise disjoint classes:

(i)

$$1 \leq u \leq \sqrt{x} \quad \text{and} \quad 1 \leq v \leq \sqrt{x},$$

(ii)

$$1 \leq u \leq \sqrt{x} \quad \text{and} \quad \sqrt{x} < v \leq x/u,$$

(iii)

$$\sqrt{x} < u \leq x \quad \text{and} \quad 1 \leq v \leq x/u.$$

The third class consists of the lattice points (u, v) such that

$$1 \leq v \leq \sqrt{x} \quad \text{and} \quad \sqrt{x} < u \leq x/v.$$

It follows from Theorem 6.9 that

$$\begin{aligned} D(x) &= \left[\sqrt{x}\right]^2 + \sum_{1 \leq u \leq \sqrt{x}} \left(\left[\frac{x}{u}\right] - \left[\sqrt{x}\right]\right) + \sum_{1 \leq v \leq \sqrt{x}} \left(\left[\frac{x}{v}\right] - \left[\sqrt{x}\right]\right) \\ &= \left[\sqrt{x}\right]^2 + 2 \sum_{1 \leq u \leq \sqrt{x}} \left(\left[\frac{x}{u}\right] - \left[\sqrt{x}\right]\right) \end{aligned}$$

$$
\begin{aligned}
&= 2\sum_{1\leq u\leq\sqrt{x}}\left[\frac{x}{u}\right]-\left[\sqrt{x}\right]^2 \\
&= 2\sum_{1\leq u\leq\sqrt{x}}\left(\frac{x}{u}-\left\{\frac{x}{u}\right\}\right)-\left(\sqrt{x}-\left\{\sqrt{x}\right\}\right)^2 \\
&= 2x\sum_{1\leq u\leq\sqrt{x}}\frac{1}{u}-2\sum_{1\leq u\leq\sqrt{x}}\left\{\frac{x}{u}\right\}-x+O(\sqrt{x}) \\
&= 2x\left(\log\sqrt{x}+\gamma+O\left(\frac{1}{\sqrt{x}}\right)\right)-x+O(\sqrt{x}) \\
&= x\log x+(2\gamma-1)x+O(\sqrt{x}).
\end{aligned}
$$

This completes the proof. □

Theorem 7.4 *For $x \geq 1$,*

$$\Delta(x)=\sum_{n\leq x}(\log n-d(n)+2\gamma)=O\left(x^{1/2}\right).$$

Proof. By Theorem 7.3 we have

$$\sum_{n\leq x}d(n)=x\log x+(2\gamma-1)x+O\left(x^{1/2}\right).$$

By Theorem 6.4 we have

$$\sum_{n\leq x}\log n=x\log x-x+O(\log x).$$

Subtracting the first equation from the second, we obtain

$$\sum_{n\leq x}(\log n-d(n)+2\gamma)=O\left(x^{1/2}\right)-2\gamma\{x\}+O(\log x)=O\left(x^{1/2}\right).$$

□

An *ordered factorization* of the positive integer n into exactly ℓ factors is an ℓ-tuple $(d_1,\ldots,d_\ell)$ such that $n=d_1\cdots d_\ell$. The divisor function $d(n)$ counts the number of ordered factorizations of n into exactly two factors, since each factorization $n=dd'$ is completely determined by the first factor d. For every positive integer ℓ, we define the arithmetic function $d_\ell(n)$ as the number of factorizations of n into exactly ℓ factors. Then $d_1(n)=1$ and $d_2(n)=d(n)$ for all n.

Theorem 7.5 *For every $\ell \geq 1$, the function $d_\ell(n)$ is multiplicative, and*

$$d_\ell(p^a) = \binom{a+\ell-1}{\ell-1}$$

for all prime powers p^a.

Proof. Let $(m,n) = 1$. For every ordered factorization of mn into ℓ factors we can construct ordered factorizations of m and n into ℓ parts, as follows. If $mn = d_1 \cdots d_\ell$ is an ordered factorization of mn into ℓ parts, then, by Exercise 20 in Section 1.4, for each $i = 1, \ldots, \ell$ there exist unique integers e_i and f_i such that e_i divides m, f_i divides n, and $d_i = e_i f_i$. Then $m = e_1 \cdots e_\ell$ and $n = f_1 \cdots f_\ell$ are ordered factorizations of m and n, respectively. This construction is reversible, and so establishes a bijection between ordered factorizations of mn and pairs of ordered factorizations of m and n. It follows that $d_\ell(mn) = d_\ell(m)d_\ell(n)$, and so the divisor function d_ℓ is multiplicative.

An ordered factorization of the prime power p^a can be written uniquely in the form $p^a = p^{b_1} \cdots p^{b_\ell}$, where $(b_1, \ldots, b_\ell)$ is an ordered ℓ-tuple of nonnegative integers such that $b_1 + \cdots + b_\ell = a$. It follows that $d_\ell(p^a)$ is exactly the number of ordered partitions of a into exactly ℓ nonnegative parts. Imagine a sequence of $a+\ell-1$ red squares. If we choose $\ell-1$ of these squares and color them blue, then the remaining a red squares are divided into exactly ℓ subsequences (possibly empty) of consecutive red squares, separated by blue squares. Every ordered partition of a into ℓ nonnegative parts can be uniquely constructed in this way, and so $d_\ell(p^a)$ is the number of ways to choose $\ell - 1$ squares from a set of $a + \ell - 1$ squares, that is,

$$d_\ell(p^a) = \binom{a+\ell-1}{\ell-1}.$$

This completes the proof. □

Theorem 7.6 *For $\ell \geq 2$,*

$$D_\ell(x) = \sum_{n \leq x} d_\ell(n) = \frac{1}{(\ell-1)!} x \log^{\ell-1} x + O\left(x \log^{\ell-2} x\right).$$

Proof. The proof is by induction on ℓ. By Theorem 7.3, $D_2(x) = x \log x + O(x)$. Now assume that the result holds for some integer $\ell \geq 2$. The notation $\sum_{d_1 \cdots d_\ell}$ means a sum over all ordered ℓ-tuples $(d_1, \ldots, d_\ell)$ of positive integers. Applying Theorem 6.7, we obtain

$$D_{\ell+1}(x) \quad = \quad \sum_{n \leq x} d_{\ell+1}(n)$$

$$\begin{aligned}
&= \sum_{n \le x} \sum_{d_1 \cdots d_{\ell+1} = n} 1 \\
&= \sum_{n \le x} \sum_{d_1 \cdots d_\ell | n} 1 \\
&= \sum_{d_1 \cdots d_\ell \le x} \left[\frac{x}{d_1 \cdots d_\ell} \right] \\
&= x \sum_{d_1 \cdots d_\ell \le x} \frac{1}{d_1 \cdots d_\ell} + O\left(\sum_{d_1 \cdots d_\ell \le x} 1 \right) \\
&= \frac{x \log^\ell x}{\ell!} + O\left(x \log^{\ell-1} x \right) + O(D_\ell(x)) \\
&= \frac{x \log^\ell x}{\ell!} + O\left(x \log^{\ell-1} x \right).
\end{aligned}$$

This completes the proof. □

Exercises

1. Compute $d(n)$ for $11 \le n \le 20$.
2. Prove that n is prime if and only if $d(n) = 2$.
3. Prove that $d(n)$ is prime if and only if $n = p^{q-1}$, where p and q are prime numbers.
4. Prove that $d(mn) \le d(m)d(n)$ for all positive integers m and n.
5. Prove that
$$\prod_{d|n} d = n^{d(n)/2}.$$
6. Prove that
$$\sum_{n \le x} d^2(n) \gg x \log^2 x.$$
Hint: Apply the Cauchy–Schwarz inequality to $D^2(x)$.
Remark. In Theorem 7.8 we obtain an asymptotic formula for $\sum_{n \le x} d^2(n)$.
7. Let $\omega(n)$ denote the number of distinct prime divisors of n, and let $\Omega(n)$ denote the total number of prime divisors of n. Prove that
$$2^{\omega(n)} \le d(n) \le 2^{\Omega(n)}.$$
Prove that $d(n) = 2^{\omega(n)}$ if and only if n is square-free.

8. Let $\delta > 0$ and $x \geq e^e$. Prove that the number of positive integers $n \leq x$ with $d(n) \geq (\log x)^{1+\delta}$ is $O(x(\log x)^{-\delta})$.

   Hint: $D(x) = O(x \log x)$.

9. Let $r > 1$ and $x \geq e^e$. Prove that the number of positive integers $n \leq x$ with $\omega(n) \geq r \log\log x$ is $O\left(x(\log x)^{1-r\log 2}\right)$.

10. Find all positive integers $k \leq 10$ such that $4k+1$ and $6k+1$ are simultaneously prime. Let $n_k = 12k+2$. Prove that if $4k+1$ and $6k+1$ are simultaneously prime, then $d(n_k) = d(n_k+1)$.

    Remark. It is an unsolved problem to determine whether there are infinitely many integers n such that $d(n) = d(n+1)$.

11. Prove that
$$d_\ell(n) = \prod_{p|n} \binom{v_p(n)+\ell-1}{\ell-1}$$
for all positive integers ℓ and n.

12. Let $\ell \geq 1$. Prove that
$$\sum_{n\leq x} \frac{d_{\ell+1}(n)}{n} = \sum_{d\leq x} \frac{1}{d} \sum_{n\leq x/d} \frac{d_\ell(n)}{n}.$$

13. Prove that
$$\sum_{n\leq x} \frac{d(n)}{n} = \frac{\log^2 x}{2} + O(\log x).$$

14. Let $0 < \alpha < 1$. Prove that
$$\sum_{n\leq x} \frac{d(n)}{n^\alpha} = \frac{x^{1-\alpha}\log x}{1-\alpha} + O(x^{1-\alpha}).$$

15. Let $\alpha > 1$. Prove that
$$\sum_{n\leq x} \frac{d(n)}{n^\alpha} = O(1).$$

7.2 A Theorem of Ramanujan

In Theorem 7.3 we computed the mean value of the divisor function $d(n)$. In this section we shall determine the mean value of the square of the divisor function. We begin with an alternative representation for $d^2(n)$.

Theorem 7.7

$$d^2(n) = \sum_{\delta^2 \mid n} \mu(\delta) d_4\left(\frac{n}{\delta^2}\right).$$

Proof. Define the arithmetic function $\tilde{\mu}$ as follows:

$$\tilde{\mu}(n) = \begin{cases} \mu(\sqrt{n}) & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

By Exercise 1, the function $\tilde{\mu}$ is multiplicative. Since the Dirichlet convolution of multiplicative functions is multiplicative (Exercise 3 in Section 6.4), the function $\tilde{\mu} * d_4$ is multiplicative, and

$$\begin{aligned} \tilde{\mu} * d_4(n) &= \sum_{d \mid n} \tilde{\mu}(d) d_4\left(\frac{n}{d}\right) \\ &= \sum_{\delta^2 \mid n} \mu(\delta) d_4\left(\frac{n}{\delta^2}\right). \end{aligned}$$

We shall prove that $\tilde{\mu} * d_4(p^a) = (a+1)^2$ for every prime power p^a. By Theorem 7.5,

$$d_4(p^a) = \binom{a+3}{3},$$

and so

$$\tilde{\mu} * d_4(p) = \sum_{\delta^2 \mid p} \mu(\delta) d_4\left(\frac{p}{\delta^2}\right) = d_4(p) = \binom{4}{3} = 4.$$

If $a \geq 2$, then

$$\begin{aligned} \tilde{\mu} * d_4(p^a) &= \sum_{\delta^2 \mid p^a} \mu(\delta) d_4\left(\frac{p^a}{\delta^2}\right) \\ &= d_4\left(p^a\right) - d_4\left(p^{a-2}\right) \\ &= \binom{a+3}{3} - \binom{a+1}{3} \\ &= (a+1)^2. \end{aligned}$$

Since $d(p^a) = a + 1$, it follows that

$$d^2(p^a) = (a+1)^2 = \tilde{\mu} * d_4(p^a)$$

for all prime powers p^a. The functions d^2 and $\tilde{\mu} * d_4$ are both multiplicative. Since multiplicative functions are completely determined by their values on prime powers (Exercise 2 in Section 6.4), it follows that

$$d^2(n) = \tilde{\mu} * d_4(n)$$

for all positive integers n. □

Theorem 7.8 (Ramanujan)

$$\sum_{n\leq x} d^2(n) \sim \frac{1}{\pi^2} x(\log x)^3$$

as $x \to \infty$.

Proof. Applying Theorem 7.6 with $\ell = 4$, we obtain

$$D_4(x) = \frac{x\log^3 x}{6} + O(x\log^2 x).$$

By Theorem 7.7 we have

$$\begin{aligned}
\sum_{n\leq x} d^2(n) &= \sum_{n\leq x}\sum_{\delta^2|n} \mu(\delta) d_4\left(\frac{n}{\delta^2}\right) \\
&= \sum_{\delta^2 k\leq x} \mu(\delta) d_4(k) \\
&= \sum_{\delta\leq\sqrt{x}} \mu(\delta) \sum_{k\leq x/\delta^2} d_4(k) \\
&= \sum_{\delta\leq\sqrt{x}} \mu(\delta) D_4\left(\frac{x}{\delta^2}\right) \\
&= \sum_{\delta\leq\sqrt{x}} \mu(\delta)\left(\frac{x}{6\delta^2}\log^3\frac{x}{\delta^2} + O\left(\frac{x}{\delta^2}\log^2\frac{x}{\delta^2}\right)\right) \\
&= \frac{x}{6}\sum_{\delta\leq\sqrt{x}} \frac{\mu(\delta)}{\delta^2}\log^3\frac{x}{\delta^2} + O\left(x\sum_{\delta\leq\sqrt{x}}\frac{1}{\delta^2}\log^2\frac{x}{\delta^2}\right).
\end{aligned}$$

We estimate these sums separately. The first term is

$$\begin{aligned}
&\frac{x}{6}\sum_{\delta\leq\sqrt{x}} \frac{\mu(\delta)}{\delta^2}\log^3\frac{x}{\delta^2} \\
&= \frac{x}{6}\sum_{i=0}^{3}\binom{3}{i}(-1)^i \sum_{\delta\leq\sqrt{x}}\frac{\mu(\delta)}{\delta^2}\log^{3-i} x\log^i\delta^2 \\
&= \frac{x}{6}\log^3 x\sum_{\delta\leq\sqrt{x}}\frac{\mu(\delta)}{\delta^2} + O\left(x\log^2 x\sum_{\delta\leq\sqrt{x}}\frac{\log^3\delta}{\delta^2}\right) \\
&= \frac{x}{6}\left(\frac{6}{\pi^2} + O\left(\frac{1}{\sqrt{x}}\right)\right)\log^3 x + O\left(x\log^2 x\sum_{\delta\leq\sqrt{x}}\frac{\log^3\delta}{\delta^2}\right) \\
&= \frac{x\log^3 x}{\pi^2} + +O\left(x\log^2 x\right),
\end{aligned}$$

by Theorem 6.17. Similarly,

$$x\sum_{\delta\leq\sqrt{x}}\frac{1}{\delta^2}\log^2\frac{x}{\delta^2}\leq x\log^2 x\sum_{\delta\leq\sqrt{x}}\frac{1}{\delta^2}\ll x\log^2 x.$$

This completes the proof of Ramanujan's theorem. □

Exercise

1. Prove that the function $\tilde{\mu}$ is multiplicative.

7.3 Sums of Divisors

The arithmetic function $\sigma(n)$ is defined as the sum of the positive divisors of n. Thus,

$$\begin{array}{rclcl} \sigma(1) & = & 1 & = & 1, \\ \sigma(2) & = & 1+2 & = & 3, \\ \sigma(3) & = & 1+3 & = & 4, \\ \sigma(4) & = & 1+2+4 & = & 7, \\ \sigma(5) & = & 1+5 & = & 6, \end{array} \qquad \begin{array}{rclcr} \sigma(6) & = & 1+2+3+6 & = & 12, \\ \sigma(7) & = & 1+7 & = & 8, \\ \sigma(8) & = & 1+2+4+8 & = & 15, \\ \sigma(9) & = & 1+3+9 & = & 13, \\ \sigma(10) & = & 1+2+5+10 & = & 18. \end{array}$$

If $n\geq 2$, then $\sigma(n)\geq n+1$. We can use the standard factorization of n to compute $\sigma(n)$. We begin with an example. Consider $180=2^2 3^2 5$. Every divisor d of 180 is of the form $d=2^a 3^b 5^c$, where $0\leq a\leq 2, 0\leq b\leq 2$, and $0\leq c\leq 1$. We have

$$\begin{aligned}\sigma(180) &= \sum_{d|180} d\\ &= 1+2+3+4+5+6+9+10+12\\ &\quad +15+18+20+30+36+45+60+90+180\\ &= (1+2+4)(1+3+9)(1+5)\\ &= 546.\end{aligned}$$

We can compute $\sigma(n)$ in this way for any positive integer n. If d divides n, then

$$d=\prod_{p|n}p^{a_p},$$

where

$$0\leq a_p\leq v_p(n),$$

and

$$\begin{aligned}\sigma(n) &= \sum_{d|n} d \\ &= \prod_{p|n} \sum_{a_p=0}^{v_p(n)} p^{a_p} \\ &= \prod_{p|n} \frac{p^{v_p(n)+1}-1}{p-1}.\end{aligned}$$

This formula expresses $\sigma(n)$ in terms of the standard factorization of n.

Theorem 7.9 *The arithmetic function $\sigma(n)$ is multiplicative.*

Proof. Let m and n be relatively prime positive integers. Since no prime divides both m and n, we have

$$\begin{aligned}\sigma(mn) &= \prod_{p|mn} \frac{p^{v_p(mn)+1}-1}{p-1} \\ &= \prod_{p|m} \frac{p^{v_p(m)+1}-1}{p-1} \prod_{p|n} \frac{p^{v_p(n)+1}-1}{p-1} \\ &= \sigma(m)\sigma(n).\end{aligned}$$

This completes the proof. □

The ancient Greeks divided the positive integers into three classes, determined by the sum of the divisors of the integer. They called a number *perfect* if $\sigma(n) = 2n$. A number is called *abundant* if $\sigma(n) > 2n$. A number is called *deficient* if $\sigma(n) < 2n$. The smallest perfect numbers are

$$\begin{aligned}6 &= 2\cdot 3 &&= 2^1(2^2-1), \\ 28 &= 4\cdot 7 &&= 2^2(2^3-1), \\ 496 &= 16\cdot 31 &&= 2^4(2^5-1), \\ 8128 &= 64\cdot 127 &&= 2^6(2^7-1).\end{aligned}$$

Theorem 7.10 (Euler) *An even integer n is perfect if and only if there exist prime numbers p and q such that*

$$q = 2^p - 1$$

and

$$n = 2^{p-1}q.$$

Proof. If n is of this form, then q is odd and $2n = 2^p q$. It follows that

$$\begin{aligned}\sigma(n) &= \sigma(2^{p-1})\sigma(q) \\ &= (2^p - 1)(q+1) \\ &= 2^p q + (2^p - q - 1) \\ &= 2n,\end{aligned}$$

and so n is perfect.

Conversely, if n is an even perfect number, then $\sigma(n) = 2n$. Writing n in the form

$$n = 2^{k-1}m,$$

where m is odd and $k \geq 2$ (since n is even), we have

$$2^k m = 2n = \sigma(n) = \sigma(2^{k-1}m) = \sigma\left(2^{k-1}\right)\sigma(m) = (2^k - 1)\sigma(m).$$

Since $2^k - 1$ divides $2^k m$ and $2^k - 1$ is relatively prime to 2^k, Euclid's lemma implies that $2^k - 1$ divides m, and so

$$m = \left(2^k - 1\right)\ell$$

for some odd integer ℓ. Then

$$2^k\left(2^k - 1\right)\ell = \left(2^k - 1\right)\sigma\left(\left(2^k - 1\right)\ell\right).$$

If $\ell > 1$, then $1, \ell$, and $(2^k - 1)\ell$ are distinct divisors of $(2^k - 1)\ell$, and

$$2^k\ell = \sigma\left(\left(2^k - 1\right)\ell\right) \geq 1 + \ell + (2^k - 1)\ell = 2^k\ell + 1,$$

which is impossible. Therefore, $\ell = 1$ and

$$2^k = \sigma(2^k - 1) = 1 + (2^k - 1) + \sum_{\substack{d \mid (2^k - 1) \\ 1 < d < 2^k - 1}} d,$$

it follows that $2^k - 1$ has no proper divisors, that is, $2^k - 1$ is a prime number. If the exponent k were composite, then $k = k_1 k_2$ with $1 < k_1 \leq k_2 < k$, and

$$2^k - 1 = \left(2^{k_1}\right)^{k_2} - 1 = \left(2^{k_1} - 1\right)\left(1 + 2^{k_1} + 2^{2k_1} + \cdots + 2^{k_1(k_2 - 1)}\right)$$

would be composite, which is false. Therefore, $k = p$ is also prime, and $m = q = 2^p - 1$. This completes the proof. □

A prime number of the form $2^p - 1$ is called a *Mersenne prime.* (Exercise 5 in Section 1.5 and Exercise 9 in Section 3.4 are about Mersenne

primes.) By Theorem 7.10, every even perfect number is uniquely associated with a Mersenne prime. Only finitely many Mersenne primes have been discovered, so we know only finitely many even perfect numbers. A list of all Mersenne primes known in October, 1999, appears in the Notes at the end of Chapter 1.

It is an unsolved problem to decide whether there exist infinitely many even perfect numbers. We know almost nothing about odd perfect numbers, and it is an unsolved problem to decide whether even one odd perfect number exists.

Let

$$\sigma^*(n) = \sigma(n) - n = \sum_{\substack{d \mid n \\ d < n}} d.$$

We define $\sigma^*(0) = 0$. A pair (m, n) of positive integers is called an *amicable pair* if

$$\sigma^*(n) = m$$

and

$$\sigma^*(m) = n.$$

Equivalently, (m, n) is an amicable pair if $\sigma(m) = \sigma(n) = m + n$. For example, the pair $(220, 284)$ is amicable, since

$$\sigma^*(220) = 284$$

and

$$\sigma^*(284) = 220.$$

It is not known whether there exist infinitely many amicable pairs.

For every positive integer n and nonnegative integer k, there is an integer $S_k(n)$ obtained by iterating the function σ^* as follows:

$$\begin{aligned} S_0(n) &= n, \\ S_1(n) &= \sigma^*(n), \\ S_2(n) &= \sigma^*(S_1(n)) = \sigma^*\left(\sigma^*(n)\right), \\ &\vdots \\ S_{k+1}(n) &= \sigma^*(S_k(n)), \end{aligned}$$

for all positive integers k. The sequence $\{S_k(n)\}_{k=0}^{\infty}$ is called the *aliquot sequence* of n. Since there exist abundant, perfect, and deficient numbers, it can happen that $S_{k+1}(n) > S_k(n)$, $S_{k+1}(n) = S_k(n)$, or $S_{k+1}(n) < S_k(n)$, and so the aliquot sequence can oscillate up and down. Computations indicate, however, that for small n the aliquot sequence always becomes eventually periodic. For example, the aliquot sequence for 12 is

$$12, 16, 15, 9, 4, 3, 1, 0, 0, \ldots.$$

If n is a perfect number, then $S_k(n) = n$ for all k, and the sequence $\{S_k(n)\}_{k=0}^{\infty}$ is constant. If (m,n) is an amicable pair of integers, then

$$\begin{aligned} S_0(n) &= n, \\ S_1(n) &= m, \\ S_2(n) &= n, \\ S_3(n) &= m, \end{aligned}$$

and so on. Thus, the aliquot sequence for an integer in an amicable pair oscillates with period 2. It is an unsolved problem to determine if, for every positive integer n, the sequence $\{S_k(n)\}_{k=0}^{\infty}$ is eventually periodic. This is called the *Catalan–Dickson problem.*

There is a natural generalization of the "sum of the divisors" function. For any real or complex number α, we can define the arithmetic function

$$\sigma_\alpha(n) = \sum_{\substack{d|n \\ d\geq 1}} d^\alpha.$$

Then $\sigma_0(n)$ is the divisor function $d(n)$, and $\sigma_1(n) = \sigma(n)$. The function $\sigma_\alpha(n)$ is multiplicative for every number α (Exercise 8).

Exercises

1. Compute $\sigma(n)$ for $11 \leq n \leq 20$.

2. Prove that $(17296, 18416)$ is an amicable pair.

   Hint: $17296 = 2^4 \times 23 \times 47$ and $18416 = 2^4 \times 1151$.

3. Prove that $(9,363,584, 9,437,056)$ is an amicable pair.

   Hint: $9,363,584 = 2^7 \times 191 \times 383$ and $9,437,056 = 2^7 \times 73727$.

4. Let A be a set of positive integers, and let $A(x)$ denote the number of elements $a \in A$ such that $a \leq x$. The set A has *asymptotic density* α if $\lim_{x\to\infty} A(x)/x = \alpha$. Prove that the set of even perfect numbers has asymptotic density zero.

5. Prove that $\sigma(n) = n\sigma_{-1}(n)$ for every positive integer n.

6. Prove that
$$0 \leq \sum_{d|n} \frac{\log d}{d} \leq \sigma_{-1}(n)\log n.$$

7. Prove that for every number α,
$$\sum_{d|n} \frac{\log^\alpha d}{d} = o\left(\sigma_{-1}(n)\log^\alpha n\right).$$

Hint: Observe that for any $\varepsilon > 0$,

$$\sum_{d|n} \frac{\log^k d}{d} \leq \sum_{\substack{d|n \\ d \leq n^{\varepsilon}}} \frac{\varepsilon \log^k n}{d} + \sum_{\substack{d|n \\ d > n^{\varepsilon}}} \frac{\log^k d}{n^{\varepsilon}} \leq \varepsilon \sigma_{-1}(n) \log^k n + \frac{d(n) \log^k n}{n^{\varepsilon}}$$

and apply Theorem 7.2.

8. Prove that the function $\sigma_\alpha(n)$ is multiplicative for every real or complex number α.
9. Let $\alpha > 1$. Prove that

$$n^\alpha \leq \sigma_\alpha(n) \leq \zeta(\alpha) n^\alpha$$

for all positive integers n.

Hint: $\sum_{d|n} d^\alpha = \sum_{d|n} (n/d)^\alpha$.

10. Let $\alpha \geq 1$. Prove that

$$\frac{\sigma_\alpha(n)}{n^\alpha} < \prod_{p|n} \left(1 + \frac{2}{p^\alpha}\right)$$

for every integer $n \geq 2$.

11. Prove that

$$\liminf_{n \to \infty} \frac{\sigma(n)}{n} = 1.$$

12. Let $x \geq 2$ and $n = \prod_{p \leq x} p$. Prove that

$$\frac{\sigma(n)}{n} > \sum_{p \leq x} \frac{1}{p}.$$

Remark. Theorem 8.7 implies that $\limsup_{n \to \infty} \sigma(n)/n = \infty$.

13. Consider the numbers

$$\begin{array}{lllll} a_0 & = & 12,496 & = & 2^4 \times 11 \times 71 \\ a_1 & = & 14,288 & = & 2^4 \times 19 \times 47 \\ a_2 & = & 15,472 & = & 2^4 \times 967 \\ a_3 & = & 14,536 & = & 2^3 \times 23 \times 79 \\ a_4 & = & 14,264 & = & 2^3 \times 1783. \end{array}$$

Prove that if $r \in \{0, 1, 2, 3, 4\}$ and $k \equiv r \pmod 5$, then

$$S_k(12,496) = a_r,$$

and so the aliquot sequence for 12, 496 is periodic with period 5,

14. Compute the aliquot sequences $\{S_k(n)\}_{k=0}^{\infty}$ for $n = 28, 29, 30, 31, 32$.

7.4 Sums and Differences of Products

In this section we prove two theorems of Ingham about sums and differences of divisor functions. These results have beautiful interpretations in terms of the number of solutions of diophantine equations in positive integers.

Let $V(n)$ denote the number of representations of n as a sum of products of two positive integers. The function $V(n)$ counts the number of solutions in positive integers of the diophantine equation

$$n = ab + cd. \tag{7.1}$$

Let $cd = k$. Then $1 \leq k \leq n-1$ and $n-k = ab$. Since the number of solutions of $k = cd$ is $d(k)$ and the number of solutions of $n-k = ab$ is $d(n-k)$, it follows that the number of solutions of (7.1) with $cd = k$ is $d(k)d(n-k)$, and so

$$V(n) = \sum_{k=1}^{n-1} d(k)d(n-k).$$

Consider the diophantine equation

$$\ell = ab - cd. \tag{7.2}$$

For every positive integer k, the number of solutions of (7.2) with $cd = k$ and $ab = k + \ell$ is $d(k)d(k+\ell)$. Let $U_\ell(n)$ denote the number of solutions of (7.2) in positive integers with $cd = k \leq n$. Then

$$U_\ell(n) = \sum_{k=1}^{n} d(k)d(k+\ell).$$

We need the following lemma.

Lemma 7.1 *For every $x \geq 1$,*

$$\sum_{\substack{uv \leq x \\ (u,v)=1}} \frac{1}{uv} = \frac{3}{\pi^2}\log^2 x + O(\log x).$$

Proof. We define

$$f(x) = \sum_{\substack{uv \leq x \\ (u,v)=1}} \frac{1}{uv} \tag{7.3}$$

and

$$g(x) = \sum_{st \leq x} \frac{1}{st} = \sum_{n \leq x} \frac{d(n)}{n}.$$

If $st \leq x$ and r is a common divisor of s and t, then $r^2 \leq st \leq x$, and so $r \leq \sqrt{x}$ and

$$\begin{aligned} g(x) &= \sum_{st \leq x} \frac{1}{st} \\ &= \sum_{r \leq x^{1/2}} \sum_{\substack{st \leq x \\ (s,t)=r}} \frac{1}{st} \\ &= \sum_{r \leq x^{1/2}} \frac{1}{r^2} \sum_{\substack{uv \leq x/r^2 \\ (u,v)=1}} \frac{1}{uv} \\ &= \sum_{r \leq x^{1/2}} \frac{1}{r^2} f\left(\frac{x}{r^2}\right). \end{aligned}$$

Applying Möbius inversion (Exercise 7 of Section 6.3 with $\alpha = 2$), we obtain

$$\begin{aligned} f(x) &= \sum_{r \leq x^{1/2}} \frac{\mu(r)}{r^2} g\left(\frac{x}{r^2}\right) \\ &= \sum_{r \leq x^{1/2}} \frac{\mu(r)}{r^2} \sum_{n \leq x/r^2} \frac{d(n)}{n} \\ &= \sum_{nr^2 \leq x} \frac{\mu(r)d(n)}{nr^2} \\ &= \sum_{n \leq x} \frac{d(n)}{n} \sum_{r \leq (x/n)^{1/2}} \frac{\mu(r)}{r^2} \\ &= \sum_{n \leq x} \frac{d(n)}{n} \left(\frac{6}{\pi^2} + O\left(\left(\frac{n}{x}\right)^{1/2}\right)\right) \\ &= \frac{6}{\pi^2} \sum_{n \leq x} \frac{d(n)}{n} + O\left(\frac{1}{x^{1/2}} \sum_{n \leq x} \frac{d(n)}{n^{1/2}}\right) \end{aligned}$$

by Theorem 6.17. Since

$$\sum_{n \leq x} \frac{d(n)}{n} = \frac{\log^2 x}{2} + O(\log x)$$

and

$$\sum_{n \leq x} \frac{d(n)}{n^{1/2}} = 2x^{1/2} \log x + O(x^{1/2})$$

by Exercises 13 and 14 of Section 7.1, it follows that

$$f(x) = \frac{3}{\pi^2} \log^2 x + O\left(\log x\right).$$

This completes the proof. □

Theorem 7.11

$$V(n) = \sum_{k=1}^{n-1} d(k)d(n-k) \sim \frac{6}{\pi^2}\sigma(n)\log^2 n.$$

Proof. The arithmetic function $V(n)$ is the number of solutions of the equation $n = ab + cd$ in positive integers. If (a,b,c,d) is a solution of this equation, then

$$ac \cdot bd = \frac{(ab+cd)^2}{4} - \frac{(ab-cd)^2}{4} \leq \frac{n^2}{4},$$

and so $ac \leq n/2$ or $bd \leq n/2$. Let P denote the number of solutions with $ac \leq n/2$, let Q denote the number of solutions with $bd \leq n/2$, and let R denote the number of solutions with both $ac \leq n/2$ and $bd \leq n/2$. Since (a,b,c,d) is a solution if and only if (b,a,d,c) is a solution, it follows that $P = Q$ and

$$V(n) = P + Q - R = 2P - R.$$

We first compute P. For fixed positive integers a and c, let $\Phi(a,c,n)$ denote the number of solutions of the equation $ab + cd = n$ in positive integers b and d. Then

$$P = \sum_{ac \leq n/2} \Phi(a,c,n).$$

Let $r = (a,c)$ denote the greatest common divisor of a and c. If r does not divide n, then $\Phi(a,c,n) = 0$. Therefore, we can assume that r divides n, and there exist positive integers α, γ, and η such that $a = r\alpha, c = r\gamma, n = r\eta$, and $(\alpha,\gamma) = 1$. Moreover, $\Phi(a,c,n) = \Phi(\alpha,\gamma,\eta)$.

Since $(\alpha,\gamma) = 1$, there exist integers b_0 and d_0 such that $\alpha b_0 + \gamma d_0 = \eta$, and every solution of the equation $ab + cd = n$ is of the form $b = b_0 + \gamma h$ and $d = d_0 - \alpha h$ for some integer h. It follows that every solution of the equation $ab + cd = n$ is of the form $b = b_0 + \gamma h$ and $d = d_0 - \alpha h$ for some integer h. If $b > 0$ and $d > 0$, then

$$-\frac{b_0}{\gamma} < h < \frac{d_0}{\alpha},$$

and so

$$\Phi(a,c,n) = \Phi(\alpha,\gamma,\eta) = \frac{b_0}{\gamma} + \frac{d_0}{\alpha} + \vartheta = \frac{\alpha b_0 + \gamma d_0}{\alpha\gamma} + \vartheta = \frac{n}{r\alpha\gamma} + \vartheta,$$

where $|\vartheta| \leq 1$ (Exercise 2). We have

$$\begin{aligned}
P &= \sum_{ac \leq n/2} \Phi(a,c,n) \\
&= \sum_{r|n} \sum_{\substack{ac \leq n/2 \\ (a,c)=r}} \Phi(a,c,n) \\
&= \sum_{r|n} \sum_{\substack{\alpha\gamma \leq n/2r^2 \\ (\alpha,\gamma)=1}} \Phi(\alpha,\gamma,\eta) \\
&= \sum_{r|n} \sum_{\substack{\alpha\gamma \leq n/2r^2 \\ (\alpha,\gamma)=1}} \left(\frac{n}{r\alpha\gamma} + \vartheta \right) \\
&= n \sum_{r|n} \frac{1}{r} \sum_{\substack{\alpha\gamma \leq n/2r^2 \\ (\alpha,\gamma)=1}} \frac{1}{\alpha\gamma} + O\left(\sum_{ac \leq n/2} 1 \right) \\
&= n \sum_{r|n} \frac{1}{r} \left(\frac{3}{\pi^2} \left(\log \frac{n}{2r^2} \right)^2 + O\left(\log \frac{n}{2r^2} \right) \right) + O\left(\sum_{k \leq n/2} d(k) \right) \\
&= \frac{3n}{\pi^2} \sum_{r|n} \frac{1}{r} \left(\log \frac{n}{2r^2} \right)^2 + O\left(n\sigma_{-1}(n) \log n \right) + O\left(n \log n \right) \\
&= \frac{3n}{\pi^2} \sum_{r|n} \frac{1}{r} \left(\log \frac{n}{2r^2} \right)^2 + O\left(\sigma(n) \log n \right) \\
&= \frac{3}{\pi^2} n\sigma_{-1}(n) \log^2 n + o\left(n\sigma_{-1}(n) \log^2 n \right) + O\left(\sigma(n) \log n \right) \\
&= \frac{3}{\pi^2} \sigma(n) \log^2 n + o\left(\sigma(n) \log^2 n \right),
\end{aligned}$$

by Lemma 7.1, Theorem 7.3, and Exercises 5 and 7 in Section 7.3.

Next we compute R. For fixed integers a and c, the linear diophantine equation $ab + cd = n$ is solvable in integers if and only if n is divisible by $r = (a,c)$. Again we write $a = r\alpha, c = r\gamma$, and $n = r\eta$, where $(\alpha,\gamma) = 1$. If the integers b_0 and d_0 solve the equation $ab + cd = n$, then every solution is of the form

$$b = b_0 + h\gamma$$

and

$$d = d_0 - h\alpha$$

for some integer h.

Let a and c be positive integers with $ac \leq n/2$. Let $\Psi(a,c,n)$ denote the number of solutions of the equation $ab + cd = n$ in positive integers b and d with

$$bd \leq \frac{n}{2}.$$

Then $\Psi(a,c,n) = \Psi(\alpha,\gamma,\eta)$ counts the number of integers h such that

$$b_0 + h\gamma > 0 \qquad \text{and} \qquad d_0 - h\alpha > 0, \tag{7.4}$$

and

$$(b_0 + h\gamma)(d_0 - h\alpha) \leq \frac{n}{2}. \tag{7.5}$$

We define the rational number

$$u = \frac{a(b_0 + \gamma h)}{n}.$$

Then

$$1 - u = \frac{c(d_0 - \alpha h)}{n}.$$

Inequalities (7.4) imply that

$$0 < u < 1.$$

Inequality (7.5) implies that

$$u(1-u) \leq \frac{ac}{2n} \leq \frac{1}{4}.$$

Solving this quadratic inequality, we obtain

$$0 < u \leq \frac{1-v}{2} \tag{7.6}$$

and

$$\frac{1+v}{2} \leq u < 1, \tag{7.7}$$

where

$$v = \sqrt{1 - \frac{2ac}{n}}.$$

Note that $0 \leq v < 1$, since $0 < ac \leq n/2$. Inequality (7.6) is equivalent to

$$-\frac{b_0}{\gamma} < h \leq \frac{(1-v)nr}{2ac} - \frac{b_0}{\gamma},$$

and inequality (7.7) implies

$$0 < 1 - u \leq \frac{1-v}{2},$$

which is equivalent to

$$\frac{d_0}{\alpha} - \frac{(1-v)nr}{2ac} \leq h < \frac{d_0}{\alpha}.$$

Both of these intervals have length

$$\frac{(1-v)nr}{2ac} = \frac{(1-v^2)nr}{(1+v)2ac} \leq \frac{(1-v^2)nr}{2ac} = r.$$

It follows that if a and c are positive integers with $(a,c) \leq n/2$ and $(a,c) = r$, then

$$\Psi(a,c,n) = \frac{(1-v)nr}{ac} + O(1) \leq 2r + O(1).$$

Therefore,

$$\begin{aligned}
R &= \sum_{ac \leq n/2} \Psi(a,c,n) \\
&= \sum_{r|n} \sum_{\substack{ac \leq n/2 \\ (a,c)=r}} \Psi(a,c,n) \\
&\leq \sum_{r|n} \sum_{\substack{\alpha\gamma \leq n/(2r^2) \\ (\alpha,\gamma)=1}} (2r + O(1)) \\
&= 2\sum_{r|n} r \sum_{\alpha\gamma \leq n/(2r^2)} 1 + \sum_{ac \leq n/2} O(1) \\
&\ll \sum_{r|n} r \sum_{k \leq n/(2r^2)} d(k) + \sum_{k \leq n/2} d(k) \\
&\ll \sum_{r|n} r \left(\frac{n \log n}{r^2} \right) + n \log n \\
&\ll n\sigma_{-1}(n) \log n \\
&= \sigma(n) \log n.
\end{aligned}$$

We have

$$\begin{aligned}
V(n) &= 2P - R \\
&= \frac{6}{\pi^2}\sigma(n)\log^2 n + o(\sigma(n)\log^2 n) + O(\sigma(n)\log n) \\
&\sim \frac{6}{\pi^2}\sigma(n)\log^2 n.
\end{aligned}$$

This completes the proof. □

Theorem 7.12 *For every positive integer ℓ,*

$$U_\ell(n) = \sum_{k=1}^{n} d(k)d(k+\ell) \sim \frac{6}{\pi^2}\sigma_{-1}(\ell) n \log^2 n.$$

Proof. Let x be the geometric mean of n and $n+\ell$, that is,

$$x = \sqrt{n(n+\ell)} = n+\theta,$$

where

$$0 < \theta < \frac{\ell}{2}.$$

We have $x = O(n)$.

The function $U_\ell(n)$ counts the number of 4-tuples (a,b,c,d) of positive integers such that

$$ab - cd = \ell \qquad \text{and} \qquad cd \leq n. \tag{7.8}$$

If (a,b,c,d) satisfies (7.8), then

$$ac \cdot bd \leq n(n+\ell) = x^2,$$

and so $ac \leq x$ or $bd \leq x$. Let P be the number of solutions of (7.8) with $ac \leq x$, Q the number of solutions of (7.8) with $bd \leq x$, and R the number of solutions of (7.8) with both $ac \leq x$ and $bd \leq x$. The symmetry of equation (7.8) implies that $P = Q$, and so

$$U_\ell(n) = P + Q - R = 2P - R.$$

We shall find asymptotic formulae for P and R by the same method used in the proof of Theorem 7.11.

We first compute P. For fixed positive integers a and c, let $\Phi_\ell(a,c,n)$ denote the number of solutions of the equation $ab - cd = \ell$ in positive integers b and d with $cd \leq n$. Let $r = (a,c)$ denote the greatest common divisor of a and c. The integer r must divide ℓ, and so there exist positive integers α, γ, and λ such that $a = r\alpha, c = r\gamma, \ell = r\lambda$, and $(\alpha,\gamma) = 1$. If $cd \leq n$, then $\gamma d \leq n/r$. If $ab - cd = \ell$, then $\alpha b - \gamma d = \lambda$. If $ac \leq x$, then $\alpha\gamma \leq x/r^2$. Therefore, $\Phi_\ell(a,c,n) = \Phi_\lambda(\alpha,\gamma,n/r)$ and

$$P = \sum_{ac \leq x} \Phi_\ell(a,c,n) = \sum_{r \mid \ell} \sum_{\substack{\alpha\gamma \leq x/r^2 \\ (\alpha,\gamma)=1}} \Phi_\lambda(\alpha,\gamma,n/r).$$

Since $(\alpha,\gamma) = 1$, there exist integers b_0 and d_0 such that $\alpha b_0 - \gamma d_0 = \lambda$, and every solution of the equation $\alpha b - \gamma d = \lambda$ is of the form $b = b_0 + \gamma h$ and $d = d_0 + \alpha h$ for some integer h. It follows that every solution of the equation $ab - cd = \ell$ is of the form $b = b_0 + \gamma h$ and $d = d_0 + \alpha h$ for some integer h. If $d > 0$ and $cd \leq n$, then $b > 0$ and

$$-\frac{d_0}{\alpha} < h \leq \frac{n}{\alpha c} - \frac{d_0}{\alpha} = \frac{n}{r\alpha\gamma} - \frac{d_0}{\alpha}. \tag{7.9}$$

Conversely, if h satisfies (7.9), then b and d are positive integers with $cd \leq n$. Therefore,

$$\Phi_\ell(a,c,n) = \Phi_\lambda(\alpha,\gamma,n/r) = \frac{n}{r\alpha\gamma} + \theta,$$

where $|\vartheta| \leq 1$. We have

$$\begin{aligned}
P &= \sum_{r|\ell} \sum_{\substack{\alpha\gamma \leq x/r^2 \\ (\alpha,\gamma)=1}} \Phi_\lambda(\alpha, \gamma, n/r) \\
&= \sum_{r|\ell} \sum_{\substack{\alpha\gamma \leq x/r^2 \\ (\alpha,\gamma)=1}} \left(\frac{n}{r\alpha\gamma} + \vartheta \right) \\
&= n \sum_{r|\ell} \frac{1}{r} \sum_{\substack{\alpha\gamma \leq x/r^2 \\ (\alpha,\gamma)=1}} \frac{1}{\alpha\gamma} + O\left(\sum_{ac \leq x} 1 \right) \\
&= n \sum_{r|\ell} \frac{1}{r} \left(\frac{3}{\pi^2} \left(\log \frac{n}{r^2} \right)^2 + O\left(\log \frac{n}{r^2} \right) \right) + O\left(\sum_{k \leq x} d(k) \right) \\
&= \frac{3n}{\pi^2} \sum_{r|\ell} \frac{1}{r} \left(\log \frac{n}{r^2} \right)^2 + O\left(n\sigma_{-1}(n) \log n \right) + O\left(x \log x \right) \\
&= \frac{3n}{\pi^2} \sum_{r|\ell} \frac{1}{r} \left(\log \frac{n}{r^2} \right)^2 + O\left(\sigma(n) \log n \right) \\
&= \frac{3}{\pi^2} n\sigma_{-1}(n) \log^2 n + o\left(n\sigma_{-1}(n) \log^2 n \right) + O\left(\sigma(n) \log n \right) \\
&= \frac{3}{\pi^2} \sigma(n) \log^2 n + o\left(\sigma(n) \log^2 n \right),
\end{aligned}$$

by Lemma 7.1, Theorem 7.3, and Exercises 5 and 7 in Section 7.3.

Next we compute R, which is the number of solutions of (7.8) with both $ac \leq x$ and $bd \leq x$. For fixed positive integers a and c, we let $\Psi(a, c, \ell)$ denote the number of ordered pairs (b, d) of positive integers such that $ab - cd = \ell$ and

$$0 < d \leq \frac{n}{c} \qquad \text{and} \qquad bd \leq x.$$

If $r = (a, c)$, then $a = r\alpha$ and $c = r\gamma$, where α and γ are relatively prime positive integers. If r does not divide ℓ, then $\Psi(a, c, \ell) = 0$. If r does divide ℓ, then $\ell = r\lambda$ and $\Psi(a, c, \ell) = \Psi(\alpha, \gamma, \lambda)$. Since $(\alpha, \gamma) = 1$, there exist integers b_0 and d_0 such that

$$\alpha b_0 - \gamma d_0 = \lambda,$$

and every integral solution of the linear diophantine equation $\alpha b - \gamma d = \lambda$ is of the form

$$b = b_0 + \gamma h \qquad \text{and} \qquad d = d_0 + \alpha h$$

for some integer h. Every solution in integers of $ab - cd = \ell$ is of the form

$$b = b_0 + \gamma h \qquad \text{and} \qquad d = d_0 + \alpha h$$

for some integer h. The inequality $0 < cd \leq n$ implies that

$$-\frac{d_0}{\alpha} < h \leq \frac{n}{\alpha c} - \frac{d_0}{\alpha}.$$

Since

$$\frac{b_0}{\gamma} - \frac{d_0}{\alpha} = \frac{\alpha b_0 - \gamma d_0}{\alpha\gamma} = \frac{\lambda}{\alpha\gamma} > 0,$$

it follows that

$$0 < \frac{d_0}{\alpha} + h < \frac{b_0}{\gamma} + h.$$

If $bd = (b_0 + \gamma h)(d_0 + \alpha h) \leq x$, then

$$\left(\frac{d_0}{\alpha} + h\right)^2 < \left(\frac{b_0}{\gamma} + h\right)\left(\frac{d_0}{\alpha} + h\right) \leq \frac{x}{\alpha\gamma},$$

and so

$$0 < \frac{d_0}{\alpha} + h \leq \sqrt{\frac{x}{\alpha\gamma}}.$$

Therefore,

$$\Psi(a, c, \ell) \leq \sqrt{\frac{x}{\alpha\gamma}} + 1 \leq 2\sqrt{\frac{x}{\alpha\gamma}}$$

and

$$\begin{aligned}
R &= \sum_{ac \leq x} \Psi(a, c, \ell) \\
&= \sum_{r|\ell} \sum_{\substack{ac \leq x \\ (a,c)=r}} \Psi(a, c, \ell) \\
&\leq 2\sum_{r|\ell} \sum_{\substack{\alpha\gamma \leq x/r^2 \\ (\alpha,\gamma)=1}} \sqrt{\frac{x}{\alpha\gamma}} \\
&\leq 2\sqrt{x}\sum_{r|\ell} \sum_{\alpha\gamma \leq x/r^2} \sqrt{\frac{1}{\alpha\gamma}} \\
&= 2\sqrt{x}\sum_{r|\ell} \sum_{n \leq x/r^2} \frac{d(n)}{\sqrt{n}} \\
&\ll \sqrt{x}\sum_{r|\ell} \sqrt{\frac{x}{r^2}} \log \frac{x}{r^2} \\
&\ll x \log x,
\end{aligned}$$

by Exercise 14 in Section 7.1. This completes the proof. □

Exercises

1. Prove that the diophantine equation (7.2) has infinitely many solutions in positive integers.

2. Let x and y be real numbers with $x < y$. Prove that the number of integers in the open interval (x, y) is $y - x + \theta$, where $|\theta| \leq 1$.

7.5 Sets of Multiples

Let A be a nonempty set of positive integers. The *set of multiples* $M(A)$ consists of all positive multiples of elements of A, that is,

$$M(A) = \{ma : a \in A \text{ and } m \in \mathbf{N}\}.$$

The set B is called a *set of multiples* if $B = M(A)$ for some set A. For example, if $A = \{2\}$, then $M(A)$ is the set of positive even integers. If $\mathbf{P}$ is the set of prime numbers, then $M(\mathbf{P})$ is the set of all integers $n > 1$.

A nonempty set A of positive integers is called *primitive* if no element of A divides another element of A, that is, if $a, a' \in A$ and a divides a', then $a = a'$. If A_1 and A_2 are nonempty sets of positive integers and A_1 is a subset of A_2, then $M(A_1)$ is a subset of $M(A_2)$. If A_2 is primitive and A_1 is a proper subset of A_2, then, by Exercise 4, $M(A_1)$ is a proper subset of $M(A_2)$.

We shall prove that if B is a set of multiples, then there exists a unique primitive set A^* such that $B = M(A^*)$.

Lemma 7.2 *Let A be a nonempty set of positive integers, and let A^* be the subset of A consisting of all integers $a \in A$ not divisible by any other element of A. Then A^* is a primitive set, and*

$$M(A) = M(A^*).$$

Proof. The primitivity of the set A^* follows immediately from the definition.

If $b \in M(A)$, then b is a multiple of a for some integer $a \in A$. If $a \notin A^*$, then a has a proper divisor that belongs to A. Let a' be the smallest element of A that divides a. Then $a' \in A^*$, and b is a multiple of a'. This completes the proof. □

Lemma 7.3 *If A_1 and A_2 are nonempty sets of positive integers such that $M(A_1) = M(A_2)$, then $M(A_1 \cap A_2) = M(A_1)$.*

Proof. By Exercise 4, $M(A_1 \cap A_2)$ is a subset of $M(A_1)$. If $M(A_1 \cap A_2)$ is a proper subset of $M(A_1)$, then there exists a smallest integer $b \in M(A_1) \setminus M(A_1 \cap A_2)$. Since $b \in M(A_1) = M(A_2)$, we have

$$b = m_1 a_1 = m_2 a_2$$

for positive integers m_1, m_2, a_1, a_2 with $a_1 \in A_1$, $a_2 \in A_2$. Moreover, $a_1 \neq a_2$ since $b \notin M(A_1 \cap A_2)$. Suppose $a_1 < a_2$. Since $a_1 \in M(A_1)$ and

$$a_1 < a_2 \leq m_2 a_2 = b,$$

the minimality of b implies that $a \in M(A_1 \cap A_2)$. Then $a_1 = ma$ for some $a \in A_1 \cap A_2$, and so $b = m_1 a_1 = m_1 m a \in M(A_1 \cap A_2)$, which is absurd. It follows that $M(A_1) = M(A_1 \cap A_2)$. □

Theorem 7.13 *Let B be a set of multiples. There exists a unique primitive set A^* such that $B = M(A^*)$.*

Proof. Let $B = M(A)$ for some set A, and let A^* be the primitive subset of A constructed in Lemma 7.2. Then $B = M(A^*)$. Let A' be any set of positive integers such that $B = M(A')$. By Lemma 7.3,

$$B = M(A') = M(A' \cap A^*) = M(A^*).$$

Since $A' \cap A^*$ is a subset of A^*, it follows from Exercise 4 that $A' \cap A^* = A^*$. Thus, A^* is a subset of every set A' such that $M(A') = B$, and so A^* is the primitive set uniquely defined by

$$A^* = \bigcap_{\substack{A' \subseteq \mathbf{N} \\ M(A') = B}} A'.$$

This completes the proof. □

Let A be a set of integers. The *counting function* $A(x)$ of the set A counts the number of positive elements of A not exceeding x, that is,

$$A(x) = \sum_{\substack{a \in A \\ 1 \leq a \leq x}} 1.$$

The *lower asymptotic density* of A is

$$d_L(A) = \liminf_{x \to \infty} \frac{A(x)}{x}.$$

The *upper asymptotic density* of A is

$$d_U(A) = \limsup_{x \to \infty} \frac{A(x)}{x}.$$

The set A has *asymptotic density* $d(A) = \alpha$ if $d_L(A) = d_U(A) = \alpha$, or, equivalently,

$$d(A) = \lim_{x\to\infty} \frac{A(x)}{x}.$$

The set of multiples of a finite set of positive integers always has an asymptotic density (Exercise 6), but it is possible to construct an infinite set A such that $M(A)$ does not have an asymptotic density. The following result gives a sufficient condition for the set of multiples of an infinite set to have asymptotic density.

Theorem 7.14 *If A is an infinite set of positive integers such that*

$$\sum_{a\in A} \frac{1}{a} < \infty,$$

then the set of multiples of A has an asymptotic density.

Proof. Let $A = \{a_i\}_{i=1}^{\infty}$, where $a_1 < a_2 < \cdots$, and let $B = M(A)$. For every positive integer k, let B_k denote the set of all positive integers that are divisible by a_k but not divisible by a_i for all $i < k$. The sets B_k are pairwise disjoint, and $B = \cup_{k=1}^{\infty} B_k$. It follows that

$$B(x) = \sum_{k=1}^{\infty} B_k(x)$$

and

$$\frac{B(x)}{x} = \sum_{k=1}^{\infty} \frac{B_k(x)}{x}$$

for all $x \geq 1$. There are $[x/a_k]$ positive integers not exceeding x that are divisible by a_k, and so

$$0 \leq B_k(x) \leq \left[\frac{x}{a_k}\right] \leq \frac{x}{a_k}.$$

Equivalently,

$$0 \leq \frac{B_k(x)}{x} \leq \frac{1}{a_k}$$

for all $x > 0$. Let $\varepsilon > 0$, and choose $K_1 = K_1(\varepsilon)$ such that

$$\sum_{k=K_1+1}^{\infty} \frac{1}{a_k} < \varepsilon.$$

Then

$$0 \leq \frac{B(x)}{x} - \sum_{k=1}^{K_1} \frac{B_k(x)}{x} = \sum_{k=K_1+1}^{\infty} \frac{B_k(x)}{x} \leq \sum_{k=K_1+1}^{\infty} \frac{1}{a_k} < \varepsilon.$$

By Exercise 8, the set B_k has an asymptotic density, that is, there exists a number $\beta_k \geq 0$ such that

$$d(B_k) = \lim_{x\to\infty} \frac{B_k(x)}{x} = \beta_k.$$

Moreover, $\beta_1 = d(B_1) = 1/a_1 > 0$. For every positive integer ℓ, the density of the set of integers divisible by at least one of the integers $a_1, \ldots, a_\ell$ is $\beta_1 + \cdots + \beta_\ell$, and so

$$0 < \sum_{k=1}^{\ell} \beta_k \leq 1.$$

Therefore, the infinite series $\sum_{k=1}^{\infty} \beta_k$ converges to some number $\beta > 0$. We shall prove that the set of multiples $M(A)$ has density β, that is,

$$\lim_{x\to\infty} \frac{B(x)}{x} = \beta.$$

For every $\varepsilon > 0$ there exists an integer $K_2 = K_2(\varepsilon)$ such that

$$\sum_{k=K_2+1}^{\infty} \beta_k < \varepsilon.$$

Let $K = \max\{K_1, K_2\}$. We can choose a number $x_0 = x_0(\varepsilon)$ such that

$$\left| \frac{B_k(x)}{x} - \beta_k \right| < \frac{\varepsilon}{K}$$

for all $x \geq x_0$ and $k = 1, \ldots, K$. Then

$$\begin{aligned}
\left| \frac{B(x)}{x} - \beta \right| &= \left| \sum_{k=1}^{\infty} \frac{B_k(x)}{x} - \beta \right| \\
&< \left| \sum_{k=1}^{K} \frac{B_k(x)}{x} - \sum_{k=1}^{K} \beta_k \right| + 2\varepsilon \\
&\leq \sum_{k=1}^{K} \left| \frac{B_k(x)}{x} - \beta_k \right| + 2\varepsilon \\
&< 3\varepsilon.
\end{aligned}$$

This completes the proof. □

The following result will be used in Section 7.6 to prove that the set of abundant numbers has an asymptotic density.

Theorem 7.15 *If A is an infinite set of integers with counting function*

$$A(x) = O\left(\frac{x}{\log^2 x}\right)$$

for $x \geq 2$, then the set of multiples $M(A)$ has an asymptotic density.

Proof. By Theorem 6.10, the infinite series $\sum_{a \in A} a^{-1}$ converges. It follows from Theorem 7.14 that the set of multiples $M(A)$ has an asymptotic density. □

Exercises

1. Prove that if $1 \in A$, then $M(A) = \mathbf{N}$.

2. For every positive integer n, prove that the set $\{n+1, n+2, \ldots, 2n\}$ is primitive.

3. Let $\Omega(n)$ denote the total number of prime factors of n. For every $r \geq 1$, prove that the set $\{n \geq 1 : \Omega(n) = r\}$ is primitive.

4. Prove that if A_1 and A_2 are nonempty sets of positive integers and $A_1 \subseteq A_2$, then $M(A_1) \subseteq M(A_2)$. Prove that if A_2 is primitive and A_1 is a proper subset of A_2, then $M(A_1)$ is a proper subset of $M(A_2)$.

5. Prove that if A is a primitive set, then A has upper asymptotic density $d_U(A) \leq 1/2$.

   Hint: Let $A = \{a_i\}_{i=1}^{\infty}$, where $a_1 < a_2 < a_3 < \cdots$. Prove that each a_i can be written uniquely in the form $a_i = 2^{u_i} v_i$, where $u_i \geq 0$ and v_i is an odd positive integer. Prove that the numbers v_i are distinct, since the set A is primitive.

6. Let $x \geq 1$. Let $A = \{a_1, \ldots, a_k\}$ consist of k distinct positive integers. For every subset $A' = \{a_{i_1}, \ldots, a_{i_j}\} \subseteq A$, let $N(x, A')$ denote the number of integers up to x divisible by every element of A'. Prove that

   $$N(x, A') = \left[\frac{x}{\operatorname{lcm}(A')}\right],$$

   where $\operatorname{lcm}(A) = [a_{i_1}, \ldots, a_{i_j}]$ is the least common multiple of the integers in A'. Prove that the number of integers up to x that are divisible by no element of A is

   $$\sum_{j=0}^{k} (-1)^j \sum_{\substack{A' \subseteq A \\ |A'| = j}} N(x, A') = \sum_{j=0}^{k} (-1)^j \sum_{\substack{A' \subseteq A \\ |A'| = j}} \left[\frac{x}{\operatorname{lcm}(A')}\right].$$

Let $B = M(A)$ and let $B(x)$ be the counting function of B. Prove that

$$\begin{aligned} B(x) &= \sum_{j=1}^{k}(-1)^{j-1} \sum_{\substack{A' \subseteq A \\ |A'|=j}} \left[\frac{x}{\operatorname{lcm}(A')}\right] \\ &= x\sum_{j=1}^{k}(-1)^{j-1} \sum_{\substack{A' \subseteq A \\ |A'|=j}} \frac{1}{\operatorname{lcm}(A')} + O(1). \end{aligned}$$

Deduce that the set of multiples $M(A)$ has asymptotic density

$$d(M(A)) = \sum_{j=1}^{k}(-1)^{j-1} \sum_{\substack{A' \subseteq A \\ |A'|=j}} \frac{1}{\operatorname{lcm}(A')}.$$

7. Let $A = \{a_1, \ldots, a_k\}$ consist of k pairwise relatively prime positive integers. Prove that

$$d(M(A)) = 1 - \prod_{i=1}^{k}\left(1 - \frac{1}{a_i}\right).$$

8. Let $A = \{a_1, \ldots, a_k\}$ consist of k distinct positive integers, and let B_k be the set of positive integers divisible by a_k but not divisible by a_i for all $i < k$. Prove that the set B_k has an asymptotic density $d(B_k)$, and compute $d(B_k)$.

7.6 Abundant Numbers

In this section we consider the set of perfect and abundant numbers. For simplicity, we modify our previous terminology and call the elements of this set *abundant*. Now a positive integer n is abundant if $\sigma(n) \geq 2n$. By Exercise 2, if n is abundant, then every multiple of n is also abundant.

An integer n is called a *primitive abundant number* if n is abundant but no proper divisor of n is abundant, that is, $\sigma(n) \geq 2n$ but $\sigma(d) < 2d$ for every proper divisor d of n. The set of abundant numbers consists of all multiples of the primitive abundant numbers (Exercise 3). We shall prove that the set of abundant numbers possesses an asymptotic density.

An integer n will be called a *k-abundant number* if $\sigma(n) \geq kn$. Let A_k be the set of all k-abundant numbers.

A *primitive k-abundant number* is a positive integer n such that $\sigma(n) \geq kn$, but $\sigma(d) < kd$ for every proper divisor d of n. Let PA_k denote the set of primitive k-abundant numbers. Then $A_k = M(PA_k)$, that is, A_k is the set of

multiples of PA_k. We shall prove that the set A_k has an asymptotic density for every integer $k \geq 2$. By Theorem 7.15, A_k will have an asymptotic density if the counting function of the set PA_k of primitive k-abundant numbers is $O(x \log^{-2} x)$.

We begin with some lemmas about prime divisors. The first result states that it is rare for an integer to be divisible by a large prime power.

Lemma 7.4 *The number of positive integers $n \leq x$ divisible by some prime power $p^r \geq \log^4 x$ with $r \geq 2$ is $O(x \log^{-2} x)$.*

Proof. If p is a prime such that $p \geq \log^2 x$ and p^2 divides n, then n is divisible by a prime power $p^r \geq \log^4 x$ with $r \geq 2$. The number of such integers $n \leq x$ is $[x/p^2]$.

If $p < \log^2 x$, let u_p be the least integer such that $p^{u_p} \geq \log^4 x$. The number of integers $n \leq x$ divisible by a prime power $p^r \geq \log^4 x$ is $[x/p^{u_p}]$.

Let $N_1(x)$ denote the number of integers $n \leq x$ divisible by a prime power $p^r \geq \log^4 x$. Then

$$\begin{aligned} N_1(x) &\leq \sum_{p \geq \log^2 x} \left[\frac{x}{p^2}\right] + \sum_{p < \log^2 x} \left[\frac{x}{p^{u_p}}\right] \\ &\leq x \sum_{p \geq \log^2 x} \frac{1}{p^2} + \left(\frac{x}{\log^4 x}\right) \sum_{p < \log^2 x} 1 \\ &\leq x \sum_{n \geq \log^2 x} \frac{1}{n^2} + \left(\frac{x}{\log^4 x}\right) \log^2 x \\ &\ll \frac{x}{\log^2 x}. \end{aligned}$$

This completes the proof. □

The next result states that it is rare for a number to have many distinct prime divisors or to have only small prime divisors. Let $\omega(n)$ denote the number of distinct primes that divide n. Let $P(n)$ denote the greatest prime divisor of n.

Lemma 7.5 *Let $x \geq e^e$ and $y = \log \log x$. The number of positive integers $n \leq x$ such that either $\omega(n) \geq 5y$ or $P(n) \leq x^{1/(6y)}$ is $O(x \log^{-2} x)$ for all sufficiently large x.*

Proof. Let $N_2(x)$ denote the number of positive integers $n \leq x$ with $\omega(n) \geq 5y$. By Exercise 9 in Section 7.1,

$$N_2(x) \ll \frac{x}{(\log x)^{5 \log 2 - 1}} \leq \frac{x}{\log^2 x}.$$

Let p be a prime. If $p^r \leq x$, then $0 \leq r \leq \log x/\log p \leq \log x/\log 2$, and so the number of prime powers $p^r \leq x$ with $p \leq x^{1/(6y)}$ does not exceed

$$\left(1+\frac{\log x}{\log 2}\right) x^{1/(6y)} \ll x^{1/(6y)} \log x.$$

Let $N_3(x)$ denote the number of integers $n \leq x$ such that $\omega(n) < 5y$ and $P(n) \leq x^{1/(6y)}$. Then

$$N_3(x) \ll \left(x^{1/(6y)} \log x\right)^{5y} \ll \frac{x}{\log^2 x}$$

for all sufficiently large x. □

Combining Lemma 7.4 and Lemma 7.5, we obtain the following result.

Lemma 7.6 *There are only $O(x \log^{-2} x)$ integers $n \leq x$ that fail to satisfy all of the following three conditions:*

(i) If p^r divides n and $r \geq 2$, then $p^r < \log^4 x$.

(ii) $\omega(n) < 5y$.

(iii) $P(n) > x^{1/(6y)}$.

Lemma 7.7 *Let $n \leq x$ be a primitive k-abundant number satisfying conditions (i), (ii), and (iii) of Lemma 7.6. Then n is divisible by a prime p such that*

$$\log^4 x \leq p \leq x^{1/(13y)}. \tag{7.10}$$

Proof. If not, then we can write $n = ab$, where a is a product of primes less than $\log^4 x$, and b is a product of primes greater than $x^{1/(13y)}$. Since $x^{1/(13y)} < x^{1/(6y)}$, condition (iii) implies that $b > 1$.

By condition (ii), $\omega(b) \leq \omega(n) < 5y$. Then

$$\begin{aligned}
\frac{\sigma(b)}{b} &< \prod_{p|b}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right) \\
&\leq \prod_{p|b}\left(1+\frac{2}{p}\right) \\
&< \left(1+\frac{2}{x^{1/(13y)}}\right)^{\omega(b)} \\
&< \left(1+\frac{2}{x^{1/(13y)}}\right)^{5y} \\
&< 1+\frac{20y}{x^{1/(13y)}}
\end{aligned}$$

if x is sufficiently large (by Exercise 4 with $c = 2$). Every prime that divides a is less than $\log^4 x$, and, by condition (i), every prime power that divides n, and hence a, is also less than $\log^4 x$. Since $\omega(a) \leq \omega(n) < 5y$ by condition (ii), it follows that

$$1 \leq a < (\log^4 x)^{5y} = (\log x)^{20y}.$$

By condition (iii), $b > 1$, and so $a < n$. Since a is a proper divisor of the primitive k-abundant number n, we have

$$\sigma(a) < ka.$$

Since k is an integer, we have

$$\sigma(a) \leq ka - 1,$$

and so

$$\frac{\sigma(a)}{a} \leq k - \frac{1}{a} < k - \frac{1}{(\log x)^{20y}}.$$

Since $\sigma(n)$ is multiplicative and $n = ab$ with $(a, b) = 1$, we have, for x sufficiently large,

$$\begin{aligned}
\frac{\sigma(n)}{n} &= \frac{\sigma(a)}{a}\frac{\sigma(b)}{b} \\
&< \left(k - \frac{1}{(\log x)^{20y}}\right)\left(1 + \frac{20y}{x^{1/(13y)}}\right) \\
&< k + \frac{20ky}{x^{1/(13y)}} - \frac{1}{(\log x)^{20y}} \\
&< k,
\end{aligned}$$

which is impossible, since the integer n is k-abundant. Therefore, n must be divisible by a prime p in the interval (7.10). □

Lemma 7.8 *If x is sufficiently large and $n \leq x$ is a primitive k-abundant number satisfying conditions (i), (ii), and (iii) of Lemma 7.6, then*

$$k \leq \frac{\sigma(n)}{n} < k + \frac{k}{x^{1/(6y)}}.$$

Proof. By condition (iii), the integer n is divisible by a prime p such that

$$p \geq P(n) > x^{1/(6y)}.$$

Since $p^2 > x^{1/(3y)} > \log^4 x$ for x sufficiently large, condition (i) implies that p^2 does not divide n. Therefore $n = mp$, where $(m, p) = 1$ and $\sigma(m) < km$ since n is primitive k-abundant. It follows that

$$\frac{\sigma(n)}{n} = \frac{\sigma(m)}{m}\frac{\sigma(p)}{p} < k\left(1 + \frac{1}{p}\right) < k + \frac{k}{x^{1/(6y)}}.$$

This completes the proof. □

Theorem 7.16 *For every integer $k \geq 2$, let $PA_k(x)$ denote the number of primitive k-abundant numbers not exceeding x. Then*

$$PA_k(x) \ll \frac{x}{\log^2 x}$$

and the set A_k of k-abundant numbers possesses an asymptotic density

Proof. By Lemma 7.6 there are only $O(x \log^{-2} x)$ primitive k-abundant integers that fail to satisfy conditions (i), (ii), and (iii) of Lemma 7.6.

Let t be the number of primitive k-abundant integers $n \leq x$ that do satisfy these three conditions. We denote these numbers by $n_1, \ldots, n_t$. By Lemma 7.7, corresponding to each integer n_i there is a prime p_i such that p_i exactly divides n_i and

$$\log^4 x \leq p_i \leq x^{1/(13y)}.$$

Let $n_i = p_i m_i$. Then $(p_i, m_i) = 1$ and

$$1 \leq m_i \leq \frac{x}{\log^4 x}.$$

It suffices to prove that the integers m_i are distinct.

Suppose that $m_i = m_j$ for some $i \neq j$. Then $p_i \neq p_j$. Since

$$\frac{\sigma(n_i)}{n_i} = \frac{(p_i+1)}{p_i} \frac{\sigma(m_i)}{m_i}$$

and

$$\frac{\sigma(n_j)}{n_j} = \frac{(p_j+1)}{p_j} \frac{\sigma(m_i)}{m_i},$$

it follows that

$$\frac{\sigma(n_i)n_j}{n_i\sigma(n_j)} = \frac{(p_i+1)p_j}{p_i(p_j+1)}.$$

Since p_i and p_j are distinct primes, it follows that $(p_i+1)p_j \neq p_i(p_j+1)$. We can assume that $(p_i+1)p_j > p_i(p_j+1)$, and so

$$\begin{aligned}
\frac{\sigma(n_i)n_j}{n_i\sigma(n_j)} &= \frac{(p_i+1)p_j}{p_i(p_j+1)} \\
&\geq 1 + \frac{1}{p_i(p_j+1)} \\
&\geq 1 + \frac{1}{x^{1/(13y)}(x^{1/(13y)}+1)} \\
&\geq 1 + \frac{1}{2x^{2/(13y)}}.
\end{aligned}$$

By Lemma 7.8,

$$\frac{\sigma(n_i)n_j}{n_i\sigma(n_j)} < \left(k + \frac{k}{x^{1/(6y)}}\right)\frac{1}{k} < 1 + \frac{1}{x^{1/(6y)}}.$$

This is a contradiction, since

$$2x^{2/(13y)} < x^{1/(6y)}$$

for all sufficiently large x. It follows that the numbers $m_1, \ldots, m_t$ are distinct, and so $t \le x\log^{-4} x$. This completes the proof. □

Exercises

1. Prove that 120 is a 3-abundant number.
2. Prove that $\sigma(rn) > r\sigma(n)$ for every $r \ge 2$.
3. Prove that every abundant number is a multiple of a primitive abundant number. Prove that every k-abundant number is a multiple of a primitive k-abundant number.
4. Prove that for every $c > 1$ there exists a number $\delta_0(c) > 0$ such that for all $u > 0$ and $v > 0$ with $uv < \delta_0(c)$,

$$(1+u)^v < 1 + cuv.$$

7.7 Notes

Ramanujan stated Theorem 7.8 in [121]. Wilson [157] published a proof of this result. Ingham [69] proved Theorems 7.11 and 7.12. Johnson [75] generalized Theorem 7.11 to sums of any finite number of products. He proved that for any integer $s \ge 2$, the number of solutions in positive integers of the diophantine equation

$$n = x_1y_1 + \cdots + x_sy_s$$

is asymptotic to

$$\frac{d_{s-1}(n)\log^s n}{(s-1)!\zeta(s)},$$

where $\zeta(s)$ is the Riemann zeta function.

Besicovitch [10] constructed the first example of a set of multiples that does not have asymptotic density.

Theorem 7.16 on the asymptotic density of the abundant numbers was proved independently by Chowla [16], Erdős [31], and Davenport [19]. The proof in this book is due to Erdős. For refinements and generalizations of this result, see Elliott, *Probabilistic Number Theory I* [28, Theorem 5.6].

There are excellent research monographs on many of the topics in this chapter, for example, Halberstam and Roth, *Sequences* [48, Chapter 5], Hall, *Sets of Multiples* [49], and Hall and Tenenbaum, *Divisors* [50]. Dickson [25, Vol. I, Chapter I] is a historical catalog of results on perfect, abundant, deficient, and amicable numbers.

8
Prime Numbers

8.1 Chebyshev's Theorems

Let $\pi(x)$ denote the number of prime numbers not exceeding x, that is,

$$\pi(x) = \sum_{p \leq x} 1$$

is the counting function for the set of primes. Euclid proved that there are infinitely many primes, or, equivalently,

$$\lim_{x \to \infty} \pi(x) = \infty.$$

A classical problem in number theory is to understand the distribution of prime numbers. This problem is still fundamentally unsolved, even though we know many beautiful results about the growth of $\pi(x)$ as x tends to infinity. In this chapter we shall show that the order of magnitude of $\pi(x)$ is $x/\log x$. In Chapter 9 we shall prove the prime number theorem, which states that $\pi(x)$ is asymptotic to $x/\log x$, that is,

$$\lim_{x \to \infty} \frac{\pi(x) \log x}{x} = 1.$$

We introduce the *Chebyshev functions*

$$\vartheta(x) = \sum_{p \leq x} \log p = \log \prod_{p \leq x} p$$

and

$$\psi(x) = \sum_{p^k \leq x} \log p.$$

For example,

$$\vartheta(10) = \log 2 + \log 3 + \log 5 + \log 7$$

and

$$\psi(10) = 3\log 2 + 2\log 3 + \log 5 + \log 7.$$

The functions $\vartheta(x)$ and $\psi(x)$ count the primes $p \leq x$ and prime powers $p^k \leq x$, respectively, with weights $\log p$. Clearly,

$$\vartheta(x) \leq \psi(x).$$

If $p^k \leq x$, then $k \leq [\log x/\log p]$, and so

$$\begin{aligned} \psi(x) &= \sum_{\substack{p^k \leq x \\ k \geq 1}} \log p = \sum_{p \leq x} \left(\sum_{\substack{p^k \leq x \\ k \geq 1}} 1 \right) \log p = \sum_{p \leq x} \left[\frac{\log x}{\log p} \right] \log p \\ &\leq \sum_{p \leq x} \log x = \pi(x) \log x. \end{aligned}$$

Chebyshev proved that the functions $\vartheta(x)$ and $\psi(x)$ have order of magnitude x and that $\pi(x)$ has order of magnitude $x/\log x$.

Before proving these theorems, we need two results about binomial coefficients. The first lemma states that for fixed n, the sequence of binomial coefficients $\binom{n}{k}$ is *unimodal* in the sense that it is increasing for $k \leq n/2$ and decreasing for $k \geq n/2$. In the second lemma we apply the binomial theorem to obtain upper and lower bounds for the *middle binomial coefficient* $\binom{2n}{n}$.

Lemma 8.1 *Let $n \geq 1$ and $1 \leq k \leq n$. Then*

$$\begin{aligned} \binom{n}{k-1} &< \binom{n}{k} \quad \text{if and only if } k < \tfrac{n+1}{2}, \\ \binom{n}{k-1} &> \binom{n}{k} \quad \text{if and only if } k > \tfrac{n+1}{2}, \\ \binom{n}{k-1} &= \binom{n}{k} \quad \text{if and only if } n \text{ is odd and } k = \tfrac{n+1}{2}. \end{aligned}$$

Proof. Consider the ratio

$$r(k) = \frac{\binom{n}{k}}{\binom{n}{k-1}} = \frac{\frac{n!}{k!(n-k)!}}{\frac{n!}{(k-1)!(n-k+1)!}} = \frac{(k-1)!(n-k+1)!}{k!(n-k)!} = \frac{n-k+1}{k}.$$

Then $r(k) > 1$ if and only if $k < (n+1)/2$, and $r(k) < 1$ if and only if $k > (n+1)/2$. □

Lemma 8.2 *For all positive integers n,*

$$\frac{2^{2n}}{2n} \leq \binom{2n}{n} < 2^{2n}.$$

Proof. By the binomial theorem,

$$2^{2n} = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{n}{k} > \binom{2n}{n}.$$

By Lemma 8.1, the middle binomial coefficient $\binom{2n}{n}$ is the largest binomial coefficient in the expansion of $(1+1)^{2n}$. Therefore,

$$\begin{aligned} 2^{2n} &= \sum_{k=0}^{2n} \binom{2n}{k} = 1 + \sum_{k=1}^{2n-1} \binom{2n}{k} + 1 \\ &\leq 2 + (2n-1)\binom{2n}{n} \\ &\leq 2n\binom{2n}{n}. \end{aligned}$$

This completes the proof. □

Theorem 8.1 *For every positive integer n,*

$$\prod_{p \leq n} p < 4^n. \tag{8.1}$$

Equivalently, for every real number $x \geq 1$

$$\vartheta(x) < x \log 4. \tag{8.2}$$

Proof. Let $m \geq 1$. We consider the binomial coefficients

$$\begin{aligned} M &= \binom{2m+1}{m} = \binom{2m+1}{m+1} \\ &= \frac{(2m+1)2m(2m-1)(2m-2)\cdots(m+2)}{m!}. \end{aligned}$$

This is an integer, since M is a binomial coefficient. Moreover,

$$\begin{aligned} 2M &= \binom{2m+1}{m} + \binom{2m+1}{m+1} \\ &< \sum_{k=0}^{2m+1} \binom{2m+1}{k} \\ &= 2^{2m+1}, \end{aligned}$$

and so

$$M < 4^m.$$

If p is a prime number such that $m + 2 \leq p \leq 2m + 1$, then p divides the product

$$(2m+1)2m(2m-1)(2m-2)\cdots(m+2),$$

but p does not divide $m!$. It follows that p divides M, and so

$$\prod_{m+2\leq p\leq 2m+1} p$$

divides M. Therefore,

$$\prod_{m+2\leq p\leq 2m+1} p \leq M < 4^m \tag{8.3}$$

for all positive integers m.

We shall prove inequality (8.1) by induction on n. This inequality holds for $n = 1$ and $n = 2$, since $1 < 4^1$ and $2 < 4^2$, respectively. Let $n \geq 3$, and assume that (8.1) holds for all positive integers $m < n$. If n is even, then

$$\prod_{p\leq n} p = \prod_{p\leq n-1} p < 4^{n-1} < 4^n.$$

If n is odd, then $n = 2m + 1$ for some $m \geq 1$, and

$$\prod_{p\leq n} p = \prod_{p\leq m+1} p \prod_{m+2\leq p\leq 2m+1} p.$$

By the induction hypothesis we have

$$\prod_{p\leq m+1} p < 4^{m+1}. \tag{8.4}$$

It follows from (8.3) and (8.4) that

$$\prod_{p\leq n} p = \prod_{p\leq m+1} p \prod_{m+2\leq p\leq 2m+1} p < 4^{m+1}4^m = 4^{2m+1} = 4^n.$$

This proves (8.1).

Inequality (8.2) follows from (8.1) as follows. If $x \geq 1$, then $n = [x] \geq 1$ and

$$\vartheta(x) = \vartheta(n) = \log \prod_{p\leq n} p < n \log 4 \leq x \log 4.$$

The proof that (8.2) implies (8.1) is similar. □

We can now prove Chebyshev's theorem that the functions $\vartheta(x), \psi(x)$, and $\pi(x)\log x$ all have order of magnitude x.

Theorem 8.2 (Chebyshev) *There exist positive constants A and B such that*

$$Ax \leq \vartheta(x) \leq \psi(x) \leq \pi(x) \log x \leq B \tag{8.5}$$

for all $x \geq 2$. Moreover,

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} = \liminf_{x\to\infty} \frac{\psi(x)}{x} = \liminf_{x\to\infty} \frac{\pi(x)\log x}{x} \geq \log 2$$

and

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} = \limsup_{x\to\infty} \frac{\psi(x)}{x} = \limsup_{x\to\infty} \frac{\pi(x)\log x}{x} \leq \log 4.$$

Proof. Theorem 8.1 gives the upper bound $\vartheta(x) < x \log 4$, and so

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} \leq \log 4.$$

We shall compute a lower bound for $\psi(x)$. Let n be a positive integer, and consider the middle binomial coefficient $N = \binom{2n}{n}$. Applying Theorem 1.12, we write N as a product of prime powers as follows:

$$N = \binom{2n}{n} = \frac{(n+1)(n+2)\cdots 2n}{n!} = \frac{(2n)!}{n!^2} = \prod_{p \leq 2n} p^{v_p((2n)!) - 2v_p(n!)},$$

where

$$v_p((2n)!) - 2v_p(n!) = \sum_{k=1}^{[\log 2n/\log p]} \left(\left[\frac{2n}{p^k}\right] - 2\left[\frac{n}{p^k}\right] \right).$$

By Exercise 7, $[2t] - 2[t] = 0$ or 1 for all real numbers t, it follows that

$$0 \leq v_p((2n)!) - 2v_p(n!) \leq \left[\frac{\log 2n}{\log p}\right].$$

By Lemma 8.2,

$$\frac{2^{2n}}{2n} \leq N = \prod_{p \leq 2n} p^{v_p((2n)!) - 2v_p(n!)} \leq \prod_{p \leq 2n} p^{\left[\frac{\log 2n}{\log p}\right]},$$

and so

$$2n \log 2 - \log 2n \leq \sum_{p \leq 2n} \left[\frac{\log 2n}{\log p}\right] \log p = \psi(2n).$$

Let $x \geq 2$ and $n = [x/2]$. Then

$$2n \leq x < 2n + 2$$

and

$$\begin{aligned} \psi(x) &\geq \psi(2n) \geq 2n \log 2 - \log 2n \\ &> (x-2)\log 2 - \log x = x \log 2 - \log x - 2 \log 2. \end{aligned}$$

Therefore,

$$\liminf_{x\to\infty} \frac{\psi(x)}{x} \geq \log 2.$$

We obtain a lower bound for $\vartheta(x)$ in terms of $\pi(x)\log x$ as follows. If

$$0 < \delta < 1,$$

then

$$\begin{aligned}
\vartheta(x) &\geq \sum_{x^{1-\delta}<p\leq x} \log p \\
&\geq \sum_{x^{1-\delta}<p\leq x} (1-\delta)\log x \\
&= (1-\delta)\left(\pi(x) - \pi(x^{1-\delta})\right)\log x \\
&\geq (1-\delta)\pi(x)\log x - x^{1-\delta}\log x,
\end{aligned}$$

and so

$$\frac{\vartheta(x)}{x} \geq \frac{(1-\delta)\pi(x)\log x}{x} - \frac{\log x}{x^{\delta}}.$$

It follows that

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} \geq (1-\delta)\liminf_{x\to\infty} \frac{\pi(x)\log x}{x}.$$

This holds for all $\delta > 0$, and so

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} \geq \liminf_{x\to\infty} \frac{\pi(x)\log x}{x}. \tag{8.6}$$

Similarly,

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} \geq \limsup_{x\to\infty} \frac{\pi(x)\log x}{x}. \tag{8.7}$$

The inequality

$$\vartheta(x) \leq \psi(x) \leq \pi(x)\log x$$

implies that

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} \leq \liminf_{x\to\infty} \frac{\psi(x)}{x} \leq \liminf_{x\to\infty} \frac{\pi(x)\log x}{x} \tag{8.8}$$

and

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} \leq \limsup_{x\to\infty} \frac{\psi(x)}{x} \leq \limsup_{x\to\infty} \frac{\pi(x)\log x}{x}. \tag{8.9}$$

Inequalities (8.6) and (8.8) give

$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} = \liminf_{x\to\infty} \frac{\psi(x)}{x} = \liminf_{x\to\infty} \frac{\pi(x)\log x}{x} \geq \log 2.$$

Combining (8.7) and (8.9), we obtain

$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} = \limsup_{x\to\infty} \frac{\psi(x)}{x} = \limsup_{x\to\infty} \frac{\pi(x)\log x}{x} \leq \log 4.$$

This completes the proof. □

Theorem 8.3 *Let p_n denote the nth prime number. There exist positive constants a and b such that*

$$an \log n \leq p_n \leq bn \log n$$

for all $n \geq 2$.

Proof. By Chebyshev's inequality (8.5), there exist positive constants A and B such that

$$Ap_n \leq \pi(p_n) \log p_n = n \log p_n \leq Bp_n.$$

Let $a = B^{-1} > 0$. Since $p_n \geq n$, we have

$$p_n \geq B^{-1} n \log p_n \geq an \log n.$$

Similarly,

$$p_n \leq A^{-1} n \log p_n.$$

For n sufficiently large,

$$\begin{aligned} \log p_n &\leq \log n + \log\log p_n - \log A \\ &\leq \log n + 2\log\log p_n \\ &\leq \log n + (1/2)\log p_n, \end{aligned}$$

and so

$$\log p_n \leq 2 \log n.$$

Therefore, there exists an integer $n_0 \geq 2$ such that

$$p_n \leq A^{-1} n \log p_n \leq 2A^{-1} n \log n$$

for all $n \geq n_0$. Since $p_n/n \log n$ is bounded for $2 \leq n \leq n_0$, there exists a constant b such that $p_n \leq bn \log n$ for all $n \geq 2$. This completes the proof. □

There is a useful notation for describing the *order of magnitude* of functions. Let f be a complex-valued function with domain D, and let g be a

function on D such that $g(x) > 0$ for all $x \in D$. The domain D can be a set of real numbers or of integers. We write

$$f = O(g)$$

or

$$f \ll g$$

if there exists a constant c such that

$$|f(x)| \leq cg(x) \qquad \text{for all } x \in D.$$

For example, Chebyshev's theorem states that

$$\vartheta(x) = O(x).$$

If $D \subseteq \mathbf{R}$ and $\limsup D = \infty$, that is, if D contains arbitrarily large real numbers, then we write

$$f = o(g)$$

if

$$\lim_{\substack{x \to \infty \\ x \in D}} \frac{f(x)}{g(x)} = 0.$$

It follows from Chebyshev's theorem that

$$\pi(x) = o(x).$$

We also denote by $O(g)$ (resp. $o(g)$) any function f such that $f = O(g)$ (resp. $f = o(g)$). For example, $e^x = 1 + O(x)$ on every interval $[1, x_0]$, $\sin x = O(x)$ for all x, and $\log x = o(x^a)$ for every $a > 0$.

We say that the function f is *asymptotic* to g, written

$$f \sim g,$$

if

$$\lim_{\substack{x \to \infty \\ x \in D}} \frac{f(x)}{g(x)} = 1.$$

The *prime number theorem* states that $\pi(x) \sim x/\log x$. Since $\lim_{x\to\infty} f(x) = a$ if and only if $\liminf_{x\to\infty} f(x) = \limsup_{x\to\infty} f(x) = a$, Theorem 8.2 implies that the following asymptotic formulae are equivalent:

$$\begin{aligned} \pi(x) &\sim \frac{x}{\log x} \\ \vartheta(x) &\sim x \\ \psi(x) &\sim x. \end{aligned}$$

Exercises

1. Compute the asymptotic density of the set of prime numbers.
2. Compute the asymptotic density of the set of prime powers.

   Hint: Let $\Pi(x)$ denote the number of prime powers $p^k \leq x$. Show that $\Pi(x) = \Pi(\sqrt{x}) + (\Pi(x) - \Pi(\sqrt{x})) \ll \pi(x)$.
3. Compute the asymptotic density of the set of integers divisible by at least two distinct primes.
4. Prove that
$$\psi(x) = \vartheta(x) + O(\sqrt{x}).$$
5. Prove that $\psi(x) = \log N$, where N is the least common multiple of the positive integers not exceeding x.
6. Prove that there exist positive real numbers α and β such that
$$n^{\alpha n} < \prod_{i=1}^{n} p_i < n^{\beta n}.$$
7. Prove that $[kt] - k[t] \in \{0, 1, \ldots, k-1\}$ for all positive integers k and real numbers t.
8. Prove that there exists a constant c such that, for all x sufficiently large, there exists a prime p such that $x < p < (1+c)x$.
9. The prime number theorem states that $\vartheta(x) \sim x$. Prove that the prime number theorem implies that for every $\delta > 0$ there is a number $x_0(\delta)$ such that, for all $x \geq x_0(\delta)$, there exists a prime p such that $x < p < (1+\delta)x$.

8.2 Mertens's Theorems

We begin by describing two arithmetic functions whose values are logarithms of primes.

We define the function $\ell(n)$ by
$$\ell(n) = \begin{cases} \log p & \text{if } n = p \text{ is a prime power,} \\ 0 & \text{otherwise.} \end{cases}$$

Chebyshev's function $\vartheta(x)$ is the sum function of the ℓ-function, since
$$\sum_{n \leq x} \ell(n) = \sum_{p \leq x} \log p = \vartheta(x).$$

The *von Mangoldt function* $\Lambda(n)$ is defined by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ is a prime power,} \\ 0 & \text{otherwise.} \end{cases}$$

Chebyshev's function $\psi(x)$ is the sum function of the von Mangoldt function, since

$$\sum_{n \leq x} \Lambda(n) = \sum_{p^k \leq x} \log p = \psi(x).$$

Moreover,

$$\sum_{d \mid n} \Lambda(d) = \log n.$$

Theorem 8.4 *For $x \geq 2$,*

$$\sum_{m \leq x} \psi\left(\frac{x}{m}\right) = \sum_{d \leq x} \Lambda(d) \left[\frac{x}{d}\right] = x \log x - x + O(\log x).$$

Proof. With $f(n) = \Lambda(n)$ in Theorem 6.15, we have

$$F(x) = \sum_{n \leq x} \Lambda(n) = \psi(x),$$

and so

$$\begin{aligned} \sum_{m \leq x} \psi\left(\frac{x}{m}\right) &= \sum_{d \leq x} \Lambda(d) \left[\frac{x}{d}\right] \\ &= \sum_{n \leq x} \sum_{d \mid n} \Lambda(d) \\ &= \sum_{n \leq x} \log n \\ &= x \log x - x + O(\log x). \end{aligned}$$

The last identity comes from Theorem 6.4. □

Theorem 8.5 (Mertens) *For $x \geq 1$,*

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1) \tag{8.10}$$

and

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1). \tag{8.11}$$

Proof. Since $\psi(x) = O(x)$ by Chebyshev's theorem, we have

$$\begin{aligned}
x\log x - x + O(\log x) &= \sum_{d\leq x} \Lambda(d)\left[\frac{x}{d}\right] \\
&= \sum_{d\leq x} \Lambda(d)\left(\frac{x}{d} - \left\{\frac{x}{d}\right\}\right) \\
&= x\sum_{d\leq x}\frac{\Lambda(d)}{d} - \sum_{d\leq x}\Lambda(d)\left\{\frac{x}{d}\right\} \\
&= x\sum_{d\leq x}\frac{\Lambda(d)}{d} + O(\psi(x)) \\
&= x\sum_{d\leq x}\frac{\Lambda(d)}{d} + O(x).
\end{aligned}$$

We obtain equation (8.10) by dividing by x.

Next, we observe that

$$\begin{aligned}
\sum_{n\leq x}\frac{\Lambda(n)}{n} - \sum_{p\leq x}\frac{\log p}{p} &= \sum_{\substack{p^k\leq x\\ k\geq 2}}\frac{\log p}{p^k} \\
&\leq \sum_{p\leq x}\log p\sum_{k=2}^{\infty}\frac{1}{p^k} \\
&\leq \sum_{p\leq x}\frac{\log p}{p(p-1)} \\
&\ll 1.
\end{aligned}$$

This proves (8.11). □

Theorem 8.6

$$\sum_{n\leq x}\frac{\vartheta(n)}{n^2} = \log x + O(1).$$

Proof. We begin with the convergent series

$$\sum_{k\leq x}\frac{\ell(k)}{k^2} \leq \sum_{k=1}^{\infty}\frac{\ell(k)}{k^2} < \sum_{k=1}^{\infty}\frac{\log k}{k^2} < \infty.$$

By Theorem 6.3 applied to the function $f(t) = 1/t^2$, we have

$$\sum_{n\leq x}\frac{\vartheta(n)}{n^2} = \sum_{n\leq x}\sum_{k\leq n}\frac{\ell(k)}{n^2}$$

$$\begin{aligned}
&= \sum_{k\leq x} \ell(k) \sum_{k\leq n\leq x} \frac{1}{n^2} \\
&= \sum_{k\leq x} \ell(k) \left(\frac{1}{k} - \frac{1}{x} + O\left(\frac{1}{k^2}\right)\right) \\
&= \sum_{k\leq x} \frac{\ell(k)}{k} - \frac{\vartheta(x)}{x} + O\left(\sum_{k\leq x} \frac{\ell(k)}{k^2}\right) \\
&= \sum_{p\leq x} \frac{\log p}{p} + O(1) \\
&= \log x + O(1),
\end{aligned}$$

by Theorem 8.5. □

Theorem 8.7 (Mertens) *There exists a constant b_1 such that*

$$\sum_{p\leq x} \frac{1}{p} = \log\log x + b_1 + O\left(\frac{1}{\log x}\right)$$

for $x \geq 2$.

Proof. We can write

$$\sum_{p\leq x} \frac{1}{p} = \sum_{p\leq x} \frac{\log p}{p} \frac{1}{\log p} = \sum_{2\leq n\leq x} f(n)g(n),$$

where

$$f(n) = \begin{cases} \frac{\log p}{p} & \text{if } n = p, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$g(t) = \frac{1}{\log t} \qquad \text{for } t > 1.$$

Let

$$F(t) = \sum_{n\leq t} f(n) = \sum_{p\leq t} \frac{\log p}{p}.$$

Then $F(t) = 0$ for $t < 2$. By Theorem 8.5,

$$F(t) = \log t + r(t), \qquad \text{where } r(t) = O(1).$$

Therefore, the integral

$$\int_2^\infty \frac{r(t)}{t(\log t)^2}\, dt$$

converges absolutely, and

$$\int_x^\infty \frac{r(t)dt}{t(\log t)^2} = O\left(\frac{1}{\log x}\right).$$

By partial summation, we obtain

$$\begin{aligned}
\sum_{p\le x} \frac{1}{p} &= \sum_{n\le x} f(n)g(n) \\
&= F(x)g(x) - \int_2^x F(t)g'(t)dt \\
&= \frac{\log x + r(x)}{\log x} + \int_2^x \frac{\log t + r(t)}{t(\log t)^2}dt \\
&= 1 + O\left(\frac{1}{\log x}\right) + \int_2^x \frac{1}{t\log t}dt + \int_2^x \frac{r(t)}{t(\log t)^2}dt \\
&= \log\log x + 1 - \log\log 2 + \int_2^\infty \frac{r(t)}{t(\log t)^2}dt \\
&\quad - \int_x^\infty \frac{r(t)}{t(\log t)^2}dt + O\left(\frac{1}{\log x}\right) \\
&= \log\log x + b_1 + O\left(\frac{1}{\log x}\right),
\end{aligned}$$

where

$$b_1 = 1 - \log\log 2 + \int_2^\infty \frac{r(t)}{t(\log t)^2}dt. \tag{8.12}$$

This completes the proof. □

Theorem 8.8 (Mertens's formula) *There exists a constant γ such that for $x \ge 2$,*

$$\prod_{p\le x}\left(1 - \frac{1}{p}\right)^{-1} = e^\gamma \log x + O(1).$$

Remark. See Nathanson [2, pp. 162–165] for a proof that γ is Euler's constant, constructed in Theorem 6.9.

Proof. We begin with two observations. First, the series $\sum_p \sum_{k=2}^\infty p^{-k}/k$ converges, since

$$\sum_p \sum_{k=2}^\infty \frac{1}{kp^k} < \sum_p \sum_{k=2}^\infty \frac{1}{p^k} = \sum_p \frac{1}{p(p-1)} < \sum_{n=2}^\infty \frac{1}{n(n-1)} < \infty.$$

Let

$$b_2 = \sum_p \sum_{k=2}^\infty \frac{1}{kp^k} > 0.$$

Second, for $x \geq 2$,

$$\begin{aligned} 0 < \sum_{p>x}\sum_{k=2}^{\infty}\frac{1}{kp^k} &< \sum_{p>x}\frac{1}{p(p-1)} < \sum_{n>x}\frac{1}{n(n-1)} \\ &= \sum_{n=[x]+1}^{\infty}\left(\frac{1}{n-1}-\frac{1}{n}\right) = \frac{1}{[x]} \\ &\leq \frac{2}{x}. \end{aligned}$$

From the Taylor series

$$-\log(1-t) = \sum_{k=1}^{\infty}\frac{t^k}{k} \qquad \text{for } |t| < 1$$

and Theorem 8.7 we obtain

$$\begin{aligned} \log\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1} &= \sum_{p\leq x}\log\left(1-\frac{1}{p}\right)^{-1} \\ &= \sum_{p\leq x}\sum_{k=1}^{\infty}\frac{1}{kp^k} \\ &= \sum_{p\leq x}\frac{1}{p}+\sum_{p\leq x}\sum_{k=2}^{\infty}\frac{1}{kp^k} \\ &= \log\log x + b_1 + O\left(\frac{1}{\log x}\right) + b_2 - \sum_{p>x}\sum_{k=2}^{\infty}\frac{1}{kp^k} \\ &= \log\log x + b_1 + b_2 + O\left(\frac{1}{\log x}\right) + O\left(\frac{1}{x}\right) \\ &= \log\log x + b_1 + b_2 + O\left(\frac{1}{\log x}\right). \end{aligned}$$

Let $\gamma = b_1 + b_2$. Then

$$\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1} = e^{\gamma}\log x \exp\left(O\left(\frac{1}{\log x}\right)\right).$$

Since $\exp(t) = 1 + O(t)$ for t in any bounded interval $[0, t_0]$, and since $O\left(1/\log x\right)$ is bounded for $x \geq 2$, we have

$$\exp\left(O\left(\frac{1}{\log x}\right)\right) = 1 + O\left(\frac{1}{\log x}\right).$$

Therefore,

$$\begin{aligned}\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1} &= e^{\gamma}\log x\exp\left(O\left(\frac{1}{\log x}\right)\right)\\ &= e^{\gamma}\log x\left(1+O\left(\frac{1}{\log x}\right)\right)\\ &= e^{\gamma}\log x+O(1).\end{aligned}$$

This is Mertens's formula. □

Exercises

1. Prove that $1*\Lambda=L$, or, equivalently,

$$\sum_{d|n}\Lambda(d)=\log n.$$

   Prove that $\Lambda=\mu*L$.

2. Prove that

$$\sum_{x<p\leq 2x}\frac{\log p}{p}=O(1).$$

3. Prove that

$$\sum_{p\leq x}\frac{\log^2 p}{p}=\frac{1}{2}\log^2 x+O(\log x).$$

   Hint: Observe that

$$\sum_{p\leq x}\frac{\log^2 p}{p}=\sum_{n\leq x}\left(\frac{\ell(n)}{n}\right)\log n,$$

   and use partial summation.

4. Prove that

$$\sum_{p\leq x}\frac{\log^k p}{p}=\frac{1}{k}\log^k x+O(\log^{k-1}x)$$

   for every positive integer k.

   Hint: Use induction on k.

5. Prove that

$$\sum_{n\leq x}\frac{\ell*\ell(n)}{n}=\sum_{pq\leq x}\frac{\log p\log q}{pq}=\frac{1}{2}\log^2 x+O(\log x).$$

Hint: Observe that

$$\sum_{pq\leq x} \frac{\log p \log q}{pq} = \sum_{p\leq x} \frac{\log p}{p} \sum_{q\leq x/p} \frac{\log q}{q},$$

and use Mertens's formula (8.11).

6. Prove that

$$\sum_{n\leq x} \frac{\ell * \ell(n)}{n \log n} = \sum_{pq\leq x} \frac{\log p \log q}{pq \log pq} = \log x + O(\log\log x).$$

Hint: Use partial summation and the previous exercise.

7. Prove that

$$\limsup_{n\to\infty} \frac{\sigma(n)}{n} = \infty.$$

Hint: Use Exercise 12 in Section 7.3.

8.3 The Number of Prime Divisors of an Integer

The arithmetic function $\omega(n)$ counts the number of distinct prime divisors of the positive integer n, that is,

$$\omega(n) = \sum_{p|n} 1.$$

We have

$$\begin{array}{rclrcl} \omega(1) & = & 0, & \omega(6) & = & 2, \\ \omega(2) & = & 1, & \omega(7) & = & 1, \\ \omega(3) & = & 1, & \omega(8) & = & 1, \\ \omega(4) & = & 1, & \omega(9) & = & 1, \\ \omega(5) & = & 1 & \omega(10) & = & 2. \end{array}$$

The arithmetic function $\Omega(n)$ counts the total number of primes whose product is n, that is,

$$\Omega(n) = \sum_{p^r\|n} r.$$

We have

$$\begin{array}{rclrcl} \Omega(1) & = & 0, & \Omega(6) & = & 2, \\ \Omega(2) & = & 1, & \Omega(7) & = & 1, \\ \Omega(3) & = & 1, & \Omega(8) & = & 3, \\ \Omega(4) & = & 2, & \Omega(9) & = & 2, \\ \Omega(5) & = & 1 & \Omega(10) & = & 2. \end{array}$$

If

$$n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$$

is the standard factorization of n as a product of powers of distinct primes, then

$$\omega(n) = k$$

and

$$\Omega(n) = r_1 + r_2 + \cdots + r_k.$$

We shall prove that almost all integers up to x have $\log\log x$ distinct prime factors. We begin with estimates for the mean value and mean-squared value of $\omega(n)$

Theorem 8.9 *For $x \geq 2$,*

$$\sum_{n \leq x} \omega(n) = x \log\log x + b_1 x + O\left(\frac{x}{\log x}\right),$$

where b_1 is the positive real number defined by (8.12).

Proof. Applying Chebyshev's theorem (Theorem 8.2) and Mertens's theorem (Theorem 8.7), we obtain

$$\begin{aligned}
\sum_{n \leq x} \omega(n) &= \sum_{n \leq x} \sum_{p \mid n} 1 = \sum_{p \leq x} \sum_{\substack{n \leq x \\ p \mid n}} 1 \\
&= \sum_{p \leq x} \left[\frac{x}{p}\right] = \sum_{p \leq x} \frac{x}{p} + O\left(\pi(x)\right) \\
&= x \sum_{p \leq x} \frac{1}{p} + O\left(\frac{x}{\log x}\right) \\
&= x\left(\log\log x + b_1 + O\left(\frac{1}{\log x}\right)\right) + O\left(\frac{x}{\log x}\right) \\
&= x \log\log x + b_1 x + O\left(\frac{x}{\log x}\right).
\end{aligned}$$

□

Theorem 8.10 *For $x \geq 2$,*

$$\sum_{n \leq x} \omega(n)^2 = x(\log\log x)^2 + O(x \log\log x).$$

Proof. We have

$$\begin{aligned}
\omega(n)^2 &= \left(\sum_{p \mid n} 1\right)^2 = \left(\sum_{p_1 \mid n} 1\right)\left(\sum_{p_2 \mid n} 1\right) \\
&= \sum_{\substack{p_1 p_2 \mid n \\ p_1 \neq p_2}} 1 + \sum_{p \mid n} 1 = \sum_{\substack{p_1 p_2 \mid n \\ p_1 \neq p_2}} 1 + \omega(n).
\end{aligned}$$

By Theorem 8.9,

$$\begin{aligned}
\sum_{n\le x}\omega(n)^2 &= \sum_{n\le x}\sum_{\substack{p_1p_2|n\\ p_1\ne p_2}} 1+\sum_{n\le x}\omega(n)\\
&= \sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\sum_{\substack{n\le x\\ p_1p_2|n}} 1 + x\log\log x + O(x)\\
&= \sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\left[\frac{x}{p_1p_2}\right] + O(x\log\log x)\\
&= \sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\frac{x}{p_1p_2} + O\left(\sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}} 1\right) + O(x\log\log x)\\
&= x\sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\frac{1}{p_1p_2} + O(x\log\log x),
\end{aligned}$$

since, by the Fundamental Theorem of Arithmetic, there are at most $2x$ ordered pairs (p_1, p_2) of distinct primes such that $p_1p_2 \le x$. From Theorem 8.7, we obtain

$$\begin{aligned}
\sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\frac{1}{p_1p_2} &\le \left(\sum_{p\le x}\frac{1}{p}\right)^2\\
&= (\log\log x + O(1))^2\\
&= (\log\log x)^2 + O(\log\log x)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{\substack{p_1p_2\le x\\ p_1\ne p_2}}\frac{1}{p_1p_2} &\ge \left(\sum_{p\le\sqrt{x}}\frac{1}{p}\right)^2 - \sum_{p\le\sqrt{x}}\frac{1}{p^2}\\
&= (\log\log\sqrt{x} + O(1))^2 + O(1)\\
&= (\log\log x)^2 + O(\log\log x).
\end{aligned}$$

Therefore,

$$\sum_{n\le x}\omega(n)^2 = x(\log\log x)^2 + O(x\log\log x).$$

This completes the proof. □

We also need the following result, which is essentially Chebyshev's inequality in probability theory.

Theorem 8.11 (Chebyshev's inequality) *Let S be a finite set of integers, and let f be a real-valued function defined on S. Let μ and t be real numbers with $t > 0$. Then the number of integers $n \in S$ such that*

$$|f(n) - \mu| \geq t$$

does not exceed

$$\frac{1}{t^2} \sum_{n \in S} (f(n) - \mu)^2.$$

Proof. If $|f(n) - \mu| \geq t$, then

$$1 \leq \frac{(f(n) - \mu)^2}{t^2}$$

and

$$\begin{aligned}
\text{card}\{n \in S : |f(n) - \mu| \geq t\} &= \sum_{\substack{n \in S \\ |f(n) - \mu| \geq t}} 1 \\
&\leq \sum_{\substack{n \in S \\ |f(n) - \mu| \geq t}} \frac{(f(n) - \mu)^2}{t^2} \\
&\leq \frac{1}{t^2} \sum_{n \in S} (f(n) - \mu)^2.
\end{aligned}$$

□

Now we prove that $\omega(n)$ has "normal order" $\log \log n$ in the sense that $\omega(n)$ is close to $\log \log n$ for almost all n.

Theorem 8.12 (Hardy–Ramanujan) *For every $\delta > 0$, the number of integers $n \leq x$ such that*

$$|\omega(n) - \log \log n| \geq (\log \log x)^{\frac{1}{2} + \delta}$$

is $o(x)$.

Proof. (Turán [143]) Let S be the set of positive integers n not exceeding x, $f(n) = \omega(n)$, and $\mu = \log \log x$. Applying Chebyshev's inequality, we see that for any $t > 0$, the number of integers $n \leq x$ such that $|\omega(n) - \log \log x| \geq t$ is at most

$$\frac{1}{t^2} \sum_{n \leq x} (\omega(n) - \log \log x)^2.$$

We use Theorem8.9 and Theorem8.10 to evaluate this sum as follows:

$$\begin{aligned}
&\sum_{n\leq x}(\omega(n)-\log\log x)^2\\
&= \sum_{n\leq x}\omega(n)^2 - 2\log\log x\sum_{n\leq x}\omega(n) + \sum_{n\leq x}(\log\log x)^2\\
&= x(\log\log x)^2 + O(x\log\log x) - 2\log\log x(x\log\log x + O(x))\\
&\quad + x(\log\log x)^2 + O((\log\log x)^2)\\
&= O(x\log\log x).
\end{aligned}$$

Let $\delta > 0$ and $t = (\log\log x)^{\frac{1}{2}+\delta} - 1$. Then

$$\begin{aligned}
t^2 &> (\log\log x)^{1+2\delta} - 2(\log\log x)^{\frac{1}{2}+\delta}\\
&= (\log\log x)^{1+\delta}\left((\log\log x)^{\delta} - 2(\log\log x)^{-1/2}\right)\\
&\geq (\log\log x)^{1+\delta}
\end{aligned}$$

for x sufficiently large. Therefore, if

$$T = \{n \in S : |\omega(n) - \log\log x| \geq (\log\log x)^{\frac{1}{2}+\delta} - 1\},$$

then

$$\begin{aligned}
|T| &\ll \frac{x\log\log x}{\left((\log\log x)^{\frac{1}{2}+\delta} - 1\right)^2}\\
&< \frac{x\log\log x}{(\log\log x)^{1+\delta}}\\
&= \frac{x}{(\log\log x)^{\delta}}\\
&= o(x).
\end{aligned}$$

Let $x > e^e$. If

$$x^{1/e} \leq n \leq x,$$

then

$$0 < \log\log x - 1 \leq \log\log n \leq \log\log x.$$

If

$$|\omega(n) - \log\log n| \geq (\log\log x)^{\frac{1}{2}+\delta},$$

then

$$\begin{aligned}
|\omega(n) - \log\log x| &\geq |\omega(n) - \log\log n| - |\log\log x - \log\log n|\\
&\geq (\log\log x)^{\frac{1}{2}+\delta} - 1\\
&= t.
\end{aligned}$$

Therefore, if

$$U = \{n \in S : |\omega(n) - \log\log n| \geq (\log\log x)^{\frac{1}{2}+\delta}\},$$

then $U \subseteq T$ and so

$$|U| \leq x^{1/e} + |T| = o(x).$$

This completes the proof. □

Exercises

1. Compute $\omega(n)$ and $\Omega(n)$ for $11 \leq n \leq 20$.

2. Prove that there exists a constant b_3 such that for $x \geq 2$,

$$\frac{1}{x}\sum_{n\leq x} \Omega(n) = \log\log x + b_3 + O\left(\frac{1}{\log x}\right).$$

8.4 Notes

There are many beautiful open problems about prime numbers. Here are some examples.

1. Do there exist infinitely many primes p of the form $p = n^2 + 1$. For example, $5 = 2^2 + 1$, $17 = 4^2 + 1$, and $101 = 10^2 + 1$. The best result is due to Iwaniec [73], who proved that there exist infinitely many integers n such that $n^2 + 1$ is either prime or the product of two primes.

2. The *twin prime conjecture* states that there exist infinitely many primes p such that $p+2$ is also prime. For example, $\{11, 13\}$, $\{29, 31\}$, and $\{101, 103\}$ are twin primes.

3. The *Goldbach conjecture* states that every even number $n \geq 4$ can be written as the sum of two primes. For example, $4 = 2 + 2$, $8 = 3 + 5$, and $100 = 17 + 83$.

4. A polynomial $f(t)$ with integer coefficients has *prime divisor* p if p divides $f(t)$ for every integer t. We say that $f(t)$ represents a prime p if there is an integer n such that $f(n) = p$. Dirichlet's theorem (Theorem 10.9) states that if m and a are relatively prime integers with $m \geq 1$, then the polynomial $f(t) = mt + a$ represents infinitely many primes. These linear polynomials are the only polynomials that are known to represent infinitely many primes.

It is conjectured that if $f(t)$ is any irreducible polynomial with integer coefficients and positive leading coefficient, and if $f(t)$ has no prime divisor, then the polynomial $f(t)$ represents infinitely many primes.

An even more general conjecture, called *Schinzel's Hypothesis H* [124, 125], states that if $f_1(t), \ldots, f_r(t)$ are irreducible polynomials with positive leading coefficients, and if the polynomial $f_1(t) \cdots f_r(t)$ has no prime divisor, then there exist infinitely many n such that the r numbers $f_1(n), \ldots, f_r(n)$ are simultaneously prime. Many classical problems are special cases of this conjecture. For example, the problem about primes of the form $n^2 + 1$ is the case $r = 1$ and $f_1(t) = t^2 + 1$. The twin prime conjecture is the case $r = 2$, $f_1(t) = t$, and $f_2(t) = t + 2$.

5. A conjecture of Schinzel and Sierpiński [125] asserts that every positive rational number x can be represented as a quotient of shifted primes, that is, $x = (p+1)/(q+1)$ for primes p and q. It is known that the set of shifted primes $\{p + 1 : p \in \mathbf{P}\}$ generates a subgroup of the multiplicative group of positive rational numbers of index at most 3 (Elliott [30]).

6. Let $f_1(t), \ldots, f_r(t)$ be irreducible polynomials with integer coefficients and positive leading coefficients. Let $g(t)$ be a polynomial with integer coefficients. Suppose that there exist infinitely many positive integers N such that $N - g(t)$ is irreducible and the product $f_1(t) \cdots f_r(t)(N - g(t))$ has no prime divisor. *Schinzel's Hypothesis H_N* asserts that if N is sufficiently large, then there exists an integer n such that $N - g(n)$ is prime and $f_i(n)$ is prime for all $i = 1, \ldots, r$. The Goldbach conjecture is the special case when N is even, $r = 1$ and $f_1(t) = g(t) = t$. Note that if N is odd, then $f_1(t)(N - g(t)) = t(N - t)$ has the prime divisor 2.

7. Do there exist arbitrarily long finite arithmetic progressions of primes? Erdős asked the following more general question: If A is an infinite set of positive integers such that the series $\sum_{a \in A} a^{-1}$ diverges, then must A contain arbitrarily long finite arithmetic progressions? If the answer is yes, this would immediately imply the existence of long arithmetic progressions of prime numbers, since $\sum_{p \in \mathbf{P}} p^{-1}$ diverges (Theorem 8.7).

All these conjectures are still open, but important techniques, especially sieve methods and the circle method, have been developed to attack them, and some deep results have been obtained. More information can be found in the following books: Halberstam and Richert's *Sieve Methods* [47], Nathanson's *Additive Number Theory: The Classical Bases* [2], and Vaughan's *The Hardy-Littlewood Method* [148].

9

The Prime Number Theorem

9.1 Generalized Von Mangoldt Functions

The function $\pi(x)$ counts the number of prime numbers not exceeding x. Euclid proved that $\lim_{x\to\infty}\pi(x)=\infty$. The *prime number theorem* (PNT), conjectured independently around 1800 by Gauss and Legendre, states that $\pi(x)$ is asymptotic to $x/\log x$, that is,

$$\lim_{x\to\infty}\frac{\pi(x)\log x}{x}=1.$$

In this chapter we shall give an elementary proof of this theorem, where "elementary" means that we do not use contour integrals, Cauchy's theorem, or other results from analytic function theory, but only basic facts about arithmetic functions and the distribution of prime numbers that we proved in Chapters 6 and 8.

Recall that the von Mangoldt function $\Lambda(n)$ is equal to $\log p$ if n is a positive power of the prime p, and 0 otherwise. Let $L(n)=\log n$. Then

$$L=1*\Lambda,$$

where $1(n)=1$ for all n. By Möbius inversion, we have

$$\Lambda=\mu*L,$$

and so

$$\Lambda(n) \quad = \quad (\mu*L)(n)$$

$$\begin{aligned} &= \sum_{d|n} \mu(d)L(n/d) \\ &= L(n)\sum_{d|n}\mu(d) - \sum_{d|n}\mu(d)L(d) \\ &= -\sum_{d|n}\mu(d)L(d). \end{aligned}$$

The divisor function $d(n)$ counts the number of positive divisors of n. Since $d = 1 * 1$, from Möbius inversion we obtain $1 = \mu * d$, and so

$$\Lambda - 1 = \mu * L - \mu * d = \mu * (L - d).$$

For every nonnegative integer r we define the *generalized von Mangoldt function* Λ_r by

$$\Lambda_r = \mu * L^r.$$

Then $\Lambda_0 = \mu * 1 = \delta$, and $\Lambda_1 = \mu * L = \Lambda$ is the usual von Mangoldt function. The elementary proof of the prime number theorem makes use of the generalized von Mangoldt function Λ_2. We have

$$\begin{array}{lcllcl} \Lambda_2(1) &=& 0, & \Lambda_2(6) &=& 2\log 2\log 3, \\ \Lambda_2(2) &=& \log^2 2, & \Lambda_2(7) &=& \log^2 7, \\ \Lambda_2(3) &=& \log^2 3, & \Lambda_2(8) &=& 5\log^2 2, \\ \Lambda_2(4) &=& 3\log^2 2, & \Lambda_2(9) &=& 3\log^2 3, \\ \Lambda_2(5) &=& \log^2 5, & \Lambda_2(10) &=& 2\log 2\log 5. \end{array}$$

Theorem 9.1 *For every positive integer n,*

$$\Lambda_2(n) = \Lambda(n)\log n + \Lambda * \Lambda(n).$$

Proof. Recall that pointwise multiplication by the logarithm function $L(n)$ is a derivation on the ring of arithmetic functions (Theorem 6.2). Multiplying the identity $L = 1 * \Lambda$ by L, we obtain

$$\begin{aligned} L^2 &= L \cdot L \\ &= L \cdot (1 * \Lambda) \\ &= 1 * (L \cdot \Lambda) + (L \cdot 1) * \Lambda \\ &= 1 * (\Lambda \cdot L) + L * \Lambda. \end{aligned}$$

Therefore,

$$\Lambda_2 = \mu * L^2 = \mu * 1 * (\Lambda \cdot L) + \mu * L * \Lambda = \Lambda \cdot L + \Lambda * \Lambda,$$

which is the formula we want. □

We can compute the function $\Lambda_2 = \mu * L^2$ explicitly. Let $\omega(n)$ denote the number of distinct prime divisors of n. If $\omega(n) = 0$, then $n = 1$ and

$$\Lambda_2(1) = \mu(1)L(1)^2 = 0.$$

If $\omega(n) = 1$, then $n = p^k$, where p is prime, k is a positive integer, and so

$$\begin{aligned}
\Lambda_2(p^k) &= \mu(1)L^2(p^k) + \mu(p)L^2\left(p^{k-1}\right) \\
&= (k \log p)^2 - ((k-1)\log p)^2 \\
&= (2k-1)\log^2 p.
\end{aligned}$$

If $\omega(n) = 2$, then $n = p^k q^\ell$, where p and q are distinct primes, k and ℓ are positive integers, and

$$\begin{aligned}
\Lambda_2(p^k q^\ell) &= \mu(1)L^2(p^k q^\ell) + \mu(p)L^2(p^{k-1}q^\ell) + \mu(q)L^2(p^k q^{\ell-1}) \\
&\quad + \mu(pq)L^2(p^{k-1}q^{\ell-1}) \\
&= L^2(p^k q^\ell) - L^2(p^{k-1}q^\ell) - L^2(p^k q^{\ell-1}) + L^2(p^{k-1}q^{\ell-1}) \\
&= 2\log p \log q.
\end{aligned}$$

Let $\omega(n) \geq 3$. If $n = dk$, then either d or k is divisible by at least two distinct primes, and so $\Lambda(d)\Lambda(k) = 0$. Moreover, $\Lambda(n) = 0$. Applying Theorem 9.1, we have

$$\Lambda_2(n) = L(n)\Lambda(n) + \sum_{dk=n} \Lambda(d)\Lambda(k) = 0.$$

The *support* of an arithmetic function $f(n)$ is the set of all positive integers n such that $f(n) \neq 0$. We have just shown that the support of $\Lambda_2(n)$ is the set of all integers n with $\omega(n) = 1$ or 2.

Exercises

1. Compute $\Lambda_2(30)$ directly from the definition $\Lambda_2(n) = \mu * L^2$.

2. Prove that
$$\Lambda * \Lambda = -\mu L * L.$$

3. Prove that
$$L^3 = L^2 * \Lambda + 2L * L\Lambda + 1 * L^2\Lambda$$
and
$$\Lambda_3 = \Lambda_2 * \Lambda + L\Lambda_2.$$
Prove that the support of Λ_3 is the set of all integers n such that $1 \leq \omega(n) \leq 3$.

4. Prove that
$$L^{r+1} = \sum_{k=0}^{r} \binom{r}{k} L^{r-k} * L^k \Lambda$$
for all $r \geq 0$.

Hint: Use $L = 1 * \Lambda$ and Exercise 6 in Section 6.1.

5. Prove that
$$\Lambda_{r+1} = L\Lambda_r + \Lambda * \Lambda_r$$
for all $r \geq 0$.

6. Let $r \geq 1$. Prove that the support of Λ_r is the set of all positive integers n such that $1 \leq \omega(n) \leq r$.

7. For a positive number x and positive integers d and n, define
$$\lambda(d) = \lambda_x(d) = \mu(d) \log^2 \frac{x}{d}$$
and
$$\theta(n) = \theta_x(n) = 1 * \lambda(n) = \sum_{d|n} \lambda(d).$$
Prove that:

(i)
$$\theta(1) = 0.$$

(ii) If $u \geq 1$, then
$$\theta(p^u) = \log p \log \frac{x^2}{p}.$$

(iii) If $u, v \geq 1$, then
$$\theta(p^u q^v) = 2 \log p \log q.$$

(iv) If m is the product of the distinct primes dividing n, then
$$\theta(n) = \theta(m).$$

(v) If n is square-free and p divides n, then
$$\theta_x(n) = \theta_x\left(\frac{n}{p}\right) - \theta_{x/p}\left(\frac{n}{p}\right).$$

(vi) If n is divisible by three or more primes, then
$$\theta(n) = 0.$$

Hint: Reduce to the case of square-free integers n, and use induction on the number of prime factors of n.

9.2 Selberg's Formulae

The elementary proof of the prime number theorem begins with a formula of Atle Selberg for a sum over products of primes not exceeding x. We give several versions of this formula.

Theorem 9.2 (Selberg's formula) *For $x \geq 1$, the mean value of the generalized von Mangoldt function Λ_2 is*

$$\sum_{n \leq x} \Lambda_2(n) = 2x \log x + O(x). \tag{9.1}$$

Proof. We begin with a computation that uses the estimates in Theorems 6.9, 6.11, 6.12, and 6.16.

$$\begin{aligned}
\sum_{n \leq x} \Lambda_2(n) &= \sum_{n \leq x} \mu * L^2(n) \\
&= \sum_{dk \leq x} \mu(d) \log^2 k \\
&= \sum_{d \leq x} \mu(d) \sum_{k \leq x/d} \log^2 k \\
&= \sum_{d \leq x} \mu(d) \left(\frac{x}{d} \left(\log \frac{x}{d} \right)^2 - \frac{2x}{d} \log \frac{x}{d} + \frac{2x}{d} + O\left(\log^2 \frac{x}{d} \right) \right) \\
&= x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \left(\log \frac{x}{d} - 2 \right) + 2x \sum_{d \leq x} \frac{\mu(d)}{d} + O\left(\sum_{d \leq x} \log^2 \frac{x}{d} \right) \\
&= x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \left(\log \frac{x}{d} - 2 \right) + O(x) \\
&= x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \left(\sum_{m \leq x/d} \frac{1}{m} - \gamma - 2 + O\left(\frac{d}{x} \right) \right) + O(x) \\
&= x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \sum_{m \leq x/d} \frac{1}{m} - (\gamma + 2)x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} + O(x).
\end{aligned}$$

We estimate these two sums separately. The first sum gives the main term in Selberg's formula:

$$\begin{aligned}
& x \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \sum_{m \leq x/d} \frac{1}{m} \\
&\quad = x \sum_{dm \leq x} \frac{\mu(d)}{dm} \log \frac{x}{d}
\end{aligned}$$

$$
\begin{aligned}
&= x\sum_{n\leq x}\frac{1}{n}\sum_{d|n}\mu(d)\log\frac{x}{d} \\
&= x\log x\sum_{n\leq x}\frac{1}{n}\sum_{d|n}\mu(d) - x\sum_{n\leq x}\frac{1}{n}\sum_{d|n}\mu(d)\log d \\
&= x\log x + x\sum_{n\leq x}\frac{\Lambda(n)}{n} \\
&= 2x\log x + O(x),
\end{aligned}
$$

by Mertens's formula (8.10). Finally, using Theorem 6.16, we obtain

$$
\begin{aligned}
\sum_{d\leq x}\frac{\mu(d)}{d}\log\frac{x}{d} &= \sum_{d\leq x}\frac{\mu(d)}{d}\left(\sum_{m\leq x/d}\frac{1}{m} - \gamma + O\left(\frac{d}{x}\right)\right) \\
&= \sum_{dm\leq x}\frac{\mu(d)}{dm} - \gamma\sum_{d\leq x}\frac{\mu(d)}{d} + O(1) \\
&= \sum_{n\leq x}\frac{1}{n}\sum_{d|n}\mu(d) + O(1) \\
&= O(1).
\end{aligned}
$$

This completes the proof. □

Notation. By $\sum_{pq\leq x}$ we denote the sum over all ordered pairs of primes (p,q) such that $pq \leq x$. For example,

$$
\begin{aligned}
\sum_{pq\leq 8}\log p\log q &= \log 2\log 2 + \log 2\log 3 + \log 3\log 2 \\
&= \log^2 2 + 2\log 2\log 3.
\end{aligned}
$$

In the elementary proof of the prime number theorem we shall use the following equivalent forms of Theorem 9.2.

Theorem 9.3 (Selberg's formulae) *For* $x \geq 1$,

$$\sum_{p\leq x}\log^2 p + \sum_{pq\leq x}\log p\log q = 2x\log x + O(x), \tag{9.2}$$

$$\vartheta(x)\log x + \sum_{p\leq x}\log p\ \vartheta\left(\frac{x}{p}\right) = 2x\log x + O(x), \tag{9.3}$$

$$\sum_{p\leq x}\log p + \sum_{pq\leq x}\frac{\log p\log q}{\log pq} = 2x + O\left(\frac{x}{1+\log x}\right). \tag{9.4}$$

Proof. By Theorem 9.1,

$$\sum_{n\leq x} \Lambda_2(n) = \sum_{n\leq x} \Lambda(n)\log n + \sum_{n\leq x} \Lambda * \Lambda(n).$$

We consider the last two sums separately. The first sum is

$$\sum_{n\leq x} \Lambda(n)\log n = \sum_{p\leq x} \log^2 p + \sum_{\substack{p^k\leq x \\ k\geq 2}} k\log^2 p.$$

If $p^k \leq x$ and $k \geq 2$, then $p \leq \sqrt{x}$, and so

$$\begin{aligned}
\sum_{\substack{p^k\leq x \\ k\geq 2}} k\log^2 p &= \sum_{p\leq\sqrt{x}} \log^2 p \sum_{k=2}^{\left[\frac{\log x}{\log p}\right]} k \\
&\leq \sum_{p\leq\sqrt{x}} \log^2 p \left(\frac{\log x}{\log p}\right)^2 \\
&\leq \sqrt{x}\log^2 x \\
&\ll x.
\end{aligned}$$

Therefore,

$$\sum_{n\leq x} \Lambda(n)\log n = \sum_{p\leq x} \log^2 p + O(x).$$

For the second sum, we have

$$\begin{aligned}
\sum_{n\leq x} \Lambda * \Lambda(n) &= \sum_{n\leq x}\sum_{n=uv} \Lambda(u)\Lambda(v) \\
&= \sum_{\substack{p^k q^\ell\leq x \\ k,\ell\geq 1}} \log p\log q \\
&= \sum_{pq\leq x} \log p\log q + \sum_{\substack{p^k q^\ell\leq x \\ k,\ell\geq 1 \\ k+\ell\geq 3}} \log p\log q.
\end{aligned}$$

We apply Chebyshev's theorem to estimate the remainder term.

$$\begin{aligned}
\sum_{\substack{p^k q^\ell\leq x \\ k+\ell\geq 3 \\ k,\ell\geq 1}} \log p\log q &\leq \sum_{\substack{p^k q^\ell\leq x \\ k\geq 2 \\ \ell\geq 1}} \log p\log q + \sum_{\substack{p^k q^\ell\leq x \\ \ell\geq 2 \\ k\geq 1}} \log p\log q \\
&= 2\sum_{\substack{p^k q^\ell\leq x \\ k\geq 2 \\ \ell\geq 1}} \log p\log q
\end{aligned}$$

$$\begin{aligned}
&= 2\sum_{\substack{p^k\le x\\ k\ge 2}}\log p \sum_{\substack{q^\ell\le x/p^k\\ \ell\ge 1}}\log q\\
&= 2\sum_{\substack{p^k\le x\\ k\ge 2}}\log p\ \psi\left(\frac{x}{p^k}\right)\\
&\ll \sum_{\substack{p^k\le x\\ k\ge 2}}\frac{x\log p}{p^k}\\
&\ll x\sum_{p\le x}\log p\sum_{k=2}^{\infty}\frac{1}{p^k}\\
&\ll x\sum_{p\le x}\frac{\log p}{p(p-1)}\\
&\ll x.
\end{aligned}$$

Therefore,

$$\sum_{n\le x}\Lambda*\Lambda(n) = \sum_{pq\le x}\log p\log q + O(x).$$

It follows from Theorem 9.2 that

$$\begin{aligned}
\sum_{n\le x}\Lambda_2(n) &= \sum_{n\le x}\Lambda(n)\log n + \sum_{n\le x}\Lambda*\Lambda(n)\\
&= \sum_{p\le x}\log^2 p + \sum_{pq\le x}\log p\log q + O(x)\\
&= 2x\log x + O(x).
\end{aligned}$$

This proves (9.2).

Recall the arithmetic function

$$\ell(n) = \begin{cases}\log n & \text{if } n \text{ is prime, and}\\ 0 & \text{otherwise.}\end{cases}$$

We have $\vartheta(x) = \sum_{n\le x}\ell(n)$, where $\vartheta(x)/x = O(1)$ by Chebyshev's theorem. Applying partial summation, we have

$$\begin{aligned}
\sum_{p\le x}\log^2 p &= \sum_{n\le x}\ell(n)\log n\\
&= \vartheta(x)\log x - \int_1^x\frac{\vartheta(t)}{t}dt\\
&= \vartheta(x)\log x + O(x).
\end{aligned}$$

Also,

$$\sum_{pq\le x}\log p\log q = \sum_{p\le x}\log p\sum_{q\le x/p}\log q = \sum_{p\le x}\log p\ \vartheta\left(\frac{x}{p}\right).$$

Inserting these two identities into (9.2), we obtain (9.3).

Consider the function $f(n) = \ell(n)\log n + \ell * \ell(n)$. We can restate formula (9.2) as follows:

$$\begin{aligned} F(x) &= \sum_{n\leq x} f(n) \\ &= \sum_{n\leq x} (\ell(n)\log n + \ell * \ell(n)) \\ &= \sum_{p\leq x} \log^2 p + \sum_{pq\leq x} \log p \log q \\ &= 2x\log x + O(x). \end{aligned}$$

Also, $F(x) = 0$ for $x < 2$. Applying partial summation, we obtain

$$\begin{aligned} \sum_{p\leq x} \log p + \sum_{pq\leq x} \frac{\log p \log q}{\log pq} &= \sum_{2\leq n\leq x} \frac{\ell(n)\log n + \ell * \ell(n)}{\log n} \\ &= \sum_{3/2<n\leq x} \frac{f(n)}{\log n} \\ &= \frac{F(x)}{\log x} + \int_2^x \frac{F(t)}{t\log^2 t} dt \\ &= \frac{2x\log x + O(x)}{\log x} + \int_2^x \frac{2t\log t + O(t)}{t\log^2 t} dt \\ &= 2x + O\left(\frac{x}{\log x}\right), \end{aligned}$$

by Exercise 1. If $x \geq e$, then

$$\frac{x}{\log x} \leq \frac{2x}{1+\log x},$$

and so

$$O\left(\frac{x}{\log x}\right) = O\left(\frac{x}{1+\log x}\right).$$

If $1 \leq x \leq e$, then

$$\frac{x}{1+\log x} \geq 1$$

and

$$0 \leq \sum_{p\leq x} \log p + \sum_{pq\leq x} \frac{\log p \log q}{\log pq} \leq \log 2,$$

and so

$$\left| x - \sum_{p\leq x} \log p + \sum_{pq\leq x} \frac{\log p \log q}{\log pq} \right| \ll 1 \ll \frac{x}{1+\log x}.$$

This completes the proof of (9.3). □

Exercises

1. Let $x \geq 2$ and $k \geq 1$. Use integration by parts to prove that
$$\int_2^x \frac{dt}{\log^k t} = \frac{x}{\log^k x} - \frac{2}{\log^k 2} + k\int_2^x \frac{dt}{\log^{k+1} t}.$$
Prove that
$$\int_2^x \frac{dt}{\log^{k+1} t} = O_k\left(\frac{x}{\log^{k+1} x}\right),$$
where the implied constant depends on k.

   Hint: Divide the interval of integration $[2, x]$ into two subintervals $[2, \sqrt{x}]$ and $[\sqrt{x}, x]$.

2. Let $x \geq 2$ and $n \geq 1$. The *logarithmic integral* is the function
$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}.$$
Prove that
$$\mathrm{li}(x) = \sum_{k=1}^{n} \frac{(k-1)!x}{\log^k x} + O_n\left(\frac{x}{\log^{n+1} x}\right),$$
where the implied constant depends on n.

   Prove that
$$\mathrm{li}(x) \sim \frac{x}{\log x}.$$

3. Show that formula (9.4) implies formula (9.3).

4. Define the positive real numbers A and a by
$$\limsup_{x\to\infty} \frac{\vartheta(x)}{x} = A$$
and
$$\liminf_{x\to\infty} \frac{\vartheta(x)}{x} = a.$$
Observe that $a \leq A$ and that the prime number theorem is equivalent to the statement that $A = a = 1$. Use Selberg's formula (9.3) to prove that
$$A + a \leq 2.$$

   Hint: Note that $\vartheta(x) \geq (a - \varepsilon)x$ for all x sufficiently large. Choose a sequence of real numbers x_i such that x_i goes to infinity and $\vartheta(x_i) \geq (A - \varepsilon)x_i$ for x_i sufficiently large. Use Theorem 8.5.

5. Use Selberg's formula (9.3) to prove that

$$A + a \geq 2.$$

Conclude that $A + a = 2$, and that the prime number theorem is equivalent to $A = a$.

9.3 The Elementary Proof

We define the remainder term $R(x)$ for Chebyshev's function $\vartheta(x)$ by

$$R(x) = \vartheta(x) - x.$$

We shall prove the prime number theorem in the form $\vartheta(x) \sim x$, or, equivalently, $R(x) = o(x)$. More precisely, we shall prove that there exist sequences of positive real numbers $\{\delta_m\}_{m=1}^{\infty}$ and $\{u_m\}_{m=1}^{\infty}$ such that $\lim_{m\to\infty} \delta_m = 0$ and

$$|R(x)| < \delta_m x \qquad \text{for all } x \geq u_m.$$

The argument is technically elementary, but delicate.

We need the following estimate.

Lemma 9.1 *For $x > e$,*

$$\sum_{p\leq x} \frac{\log p}{p\left(1 + \log \frac{x}{p}\right)} \ll \log\log x.$$

Proof. By Mertens's theorem (Theorem 8.5), for every positive integer j we have

$$\sum_{\frac{x}{e^j}<p\leq\frac{x}{e^{j-1}}} \frac{\log p}{p} = \left(\log\frac{x}{e^{j-1}} + O(1)\right) - \left(\log\frac{x}{e^j} + O(1)\right) = O(1).$$

Moreover, if

$$\frac{x}{e^j} < p \leq \frac{x}{e^{j-1}},$$

then

$$j \leq 1 + \log\frac{x}{p} < j + 1,$$

and so

$$\sum_{\frac{x}{e^j}<p\leq\frac{x}{e^{j-1}}} \frac{\log p}{p\left(1 + \log\frac{x}{p}\right)} \leq \frac{1}{j} \sum_{\frac{x}{e^j}<p\leq\frac{x}{e^{j-1}}} \frac{\log p}{p} \ll \frac{1}{j}.$$

Therefore,

$$\begin{aligned}
\sum_{p\leq x} \frac{\log p}{p\left(1+\log \frac{x}{p}\right)} &= \sum_{j=1}^{[\log x]+1} \sum_{\frac{x}{e^j}<p\leq \frac{x}{e^{j-1}}} \frac{\log p}{p\left(1+\log \frac{x}{p}\right)} \\
&\ll \sum_{j=1}^{[\log x]+1} \frac{1}{j} \\
&\ll \log\log x.
\end{aligned}$$

This completes the proof. □

Theorem 9.4 *For $x \geq 1$,*

$$|R(x)| \leq \frac{1}{\log x} \sum_{n\leq x} \left| R\left(\frac{x}{n}\right)\right| + O\left(\frac{x \log\log x}{\log x}\right).$$

Proof. Replacing $\vartheta(x)$ by $x + R(x)$ in Selberg's formula (9.3), we obtain

$$\begin{aligned}
2x\log x + O(x) &= \vartheta(x)\log x + \sum_{p\leq x} \log p \ \vartheta\left(\frac{x}{p}\right) \\
&= (x+R(x))\log x + \sum_{p\leq x} \log p \left(\frac{x}{p} + R\left(\frac{x}{p}\right)\right) \\
&= x\log x + R(x)\log x + x\sum_{p\leq x}\frac{\log p}{p} + \sum_{p\leq x} R\left(\frac{x}{p}\right)\log p \\
&= R(x)\log x + \sum_{p\leq x} R\left(\frac{x}{p}\right)\log p + 2x\log x + O(x).
\end{aligned}$$

This gives

$$R(x)\log x = -\sum_{p\leq x} R\left(\frac{x}{p}\right)\log p + O(x). \tag{9.5}$$

We denote prime numbers by p, q, and r. Let $p \leq x$. From (9.4) we have

$$\sum_{q\leq x/p} \log q + \sum_{qr\leq x/p} \frac{\log q \log r}{\log qr} = \frac{2x}{p} + O\left(\frac{x}{p\left(1+\log\frac{x}{p}\right)}\right).$$

Then

$$\sum_{pq\leq x} \log p \log q = \sum_{p\leq x} \log p \sum_{q\leq x/p} \log q$$

$$\begin{aligned}
&= 2x\sum_{p\le x}\frac{\log p}{p} - \sum_{pqr\le x}\frac{\log p\log q\log r}{\log qr} \\
&\quad + O\left(x\sum_{p\le x}\frac{\log p}{p\left(1+\log\frac{x}{p}\right)}\right) \\
&= 2x(\log x + O(1)) - \sum_{qr\le x}\frac{\log q\log r}{\log qr}\sum_{p\le x/qr}\log p \\
&\quad + O\left(x\sum_{p\le x}\frac{\log p}{p\left(1+\log\frac{x}{p}\right)}\right) \\
&= 2x\log x - \sum_{qr\le x}\frac{\log q\log r}{\log qr}\vartheta\left(\frac{x}{qr}\right) + O(x\log\log x),
\end{aligned}$$

where the error term comes from Lemma 9.1. Inserting this expression for $\sum_{pq\le x}\log p\log q$ into Selberg's formula (9.2), we obtain

$$\sum_{p\le x}\log^2 p = \sum_{pq\le x}\frac{\log p\log q}{\log pq}\vartheta\left(\frac{x}{pq}\right) + O(x\log\log x).$$

Therefore,

$$\vartheta(x)\log x = \sum_{pq\le x}\frac{\log p\log q}{\log pq}\vartheta\left(\frac{x}{pq}\right) + O(x\log\log x). \tag{9.6}$$

Replacing $\vartheta(x)$ by $x+R(x)$ in (9.6), we obtain

$$\begin{aligned}
(x+R(x))\log x &= \sum_{pq\le x}\frac{\log p\log q}{\log pq}\left(\frac{x}{pq}+R\left(\frac{x}{pq}\right)\right) + O(x\log\log x) \\
&= x\sum_{pq\le x}\frac{\log p\log q}{pq\log pq} + \sum_{pq\le x}\frac{\log p\log q}{\log pq}R\left(\frac{x}{pq}\right) \\
&\quad + O(x\log\log x).
\end{aligned}$$

By Exercise 6 in Section 8.2,

$$\sum_{pq\le x}\frac{\log p\log q}{pq\log pq} = \log x + O(\log\log x),$$

and so

$$R(x)\log x = \sum_{pq\le x}\frac{\log p\log q}{\log pq}R\left(\frac{x}{pq}\right) + O(x\log\log x). \tag{9.7}$$

Adding formulas (9.5) and (9.7), we obtain

$$\begin{aligned}
2|R(x)|\log x &\leq \sum_{p\leq x}\log p\left|R\left(\frac{x}{p}\right)\right| + \sum_{pq\leq x}\frac{\log p\log q}{\log pq}\left|R\left(\frac{x}{pq}\right)\right| \\
&\quad + O(x\log\log x) \\
&= \sum_{n\leq x}\ell(n)\left|R\left(\frac{x}{n}\right)\right| + \sum_{n\leq x}\frac{\ell*\ell(n)}{\log n}\left|R\left(\frac{x}{n}\right)\right| \\
&\quad + O(x\log\log x) \\
&= \sum_{n\leq x}\left(\ell(n)+\frac{\ell*\ell(n)}{\log n}\right)\left|R\left(\frac{x}{n}\right)\right| + O(x\log\log x).
\end{aligned}$$

We can write the partial summation formula (6.6) with $a = 0$ and $b = [x]$ as follows:

$$\sum_{n\leq x} f(n)g(n) = \sum_{n\leq x-1} F(n)(g(n)-g(n+1)) + F(x)g([x]).$$

Let

$$f(n) = \ell(n) + \frac{\ell*\ell(n)}{\log n}$$

and $g(n) = |R(x/n)|$. By Selberg's formula (9.4),

$$F(x) = \sum_{n\leq x} f(n) = \sum_{n\leq x}\left(\ell(n)+\frac{\ell*\ell(n)}{\log n}\right) = 2x + O\left(\frac{x}{1+\log x}\right).$$

Then

$$\begin{aligned}
&\sum_{n\leq x}\left(\ell(n)+\frac{\ell*\ell(n)}{\log n}\right)\left|R\left(\frac{x}{n}\right)\right| \\
&\quad = \sum_{n\leq x-1}\left(2n+O\left(\frac{n}{1+\log n}\right)\right)\left(\left|R\left(\frac{x}{n}\right)\right| - \left|R\left(\frac{x}{n+1}\right)\right|\right) \\
&\qquad + \left(2x+O\left(\frac{x}{1+\log x}\right)\right)\left|R\left(\frac{x}{[x]}\right)\right|.
\end{aligned}$$

We evaluate these terms separately. The main term is

$$\begin{aligned}
&2\sum_{n\leq x-1} n\left(\left|R\left(\frac{x}{n}\right)\right| - \left|R\left(\frac{x}{n+1}\right)\right|\right) \\
&\quad = 2\sum_{n\leq x-1} n\left|R\left(\frac{x}{n}\right)\right| - 2\sum_{n\leq x-1} n\left|R\left(\frac{x}{n+1}\right)\right| \\
&\quad = 2\sum_{n\leq x-1} n\left|R\left(\frac{x}{n}\right)\right| - 2\sum_{2\leq n\leq x}(n-1)\left|R\left(\frac{x}{n}\right)\right|
\end{aligned}$$

$$\begin{aligned}
&= 2\sum_{n\leq x}\left|R\left(\frac{x}{n}\right)\right| - 2[x]\left|R\left(\frac{x}{[x]}\right)\right| \\
&= 2\sum_{n\leq x}\left|R\left(\frac{x}{n}\right)\right| + O(x),
\end{aligned}$$

since $1 \leq x/[x] < 2$ for all $x \geq 1$, and so $\vartheta(x/[x]) = 0$ and $R(x/[x]) = O(1)$.

To evaluate the second term, we begin by observing that

$$\begin{aligned}
\left|R\left(\frac{x}{n}\right)\right| - \left|R\left(\frac{x}{n+1}\right)\right| &= \left|\vartheta\left(\frac{x}{n}\right) - \frac{x}{n}\right| - \left|\vartheta\left(\frac{x}{n+1}\right) - \frac{x}{n+1}\right| \\
&\leq \left|\vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right) - \left(\frac{x}{n} - \frac{x}{n+1}\right)\right| \\
&\leq \left|\vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right)\right| + \left|\frac{x}{n} - \frac{x}{n+1}\right| \\
&= \vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right) + \frac{x}{n} - \frac{x}{n+1} \\
&< \vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right) + \frac{x}{n^2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_{n\leq x-1}\left(\frac{n}{1+\log n}\right)\left(\left|R\left(\frac{x}{n}\right)\right| - \left|R\left(\frac{x}{n+1}\right)\right|\right) \\
&\leq \sum_{n\leq x-1}\left(\frac{n}{1+\log n}\right)\left(\vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right)\right) \\
&\quad + x\sum_{n\leq x-1}\frac{1}{n(1+\log n)}.
\end{aligned}$$

We have

$$\begin{aligned}
&\sum_{n\leq x-1}\left(\frac{n}{1+\log n}\right)\left(\vartheta\left(\frac{x}{n}\right) - \vartheta\left(\frac{x}{n+1}\right)\right) \\
&= \sum_{n\leq x-1}\left(\frac{n}{1+\log n}\right)\vartheta\left(\frac{x}{n}\right) - \sum_{2\leq n\leq x}\left(\frac{n-1}{1+\log(n-1)}\right)\vartheta\left(\frac{x}{n}\right) \\
&= \vartheta(x) + \sum_{2\leq n\leq x-1}\left(\frac{n}{1+\log n} - \frac{n-1}{1+\log(n-1)}\right)\vartheta\left(\frac{x}{n}\right) \\
&\leq \vartheta(x) + \sum_{2\leq n\leq x-1}\left(\frac{n}{1+\log n} - \frac{n-1}{1+\log n}\right)\vartheta\left(\frac{x}{n}\right) \\
&= \vartheta(x) + \sum_{2\leq n\leq x-1}\left(\frac{1}{1+\log n}\right)\vartheta\left(\frac{x}{n}\right)
\end{aligned}$$

$$\begin{aligned} &\ll x + x \sum_{2 \le n \le x-1} \frac{1}{n(1+\log n)} \\ &\ll x \log\log x, \end{aligned}$$

since

$$\sum_{n \le x-1} \frac{1}{n(1+\log n)} = O(\log\log x)$$

by Exercise 11 of Section 6.2.

The third term is simply

$$\left(2x + O\left(\frac{x}{1+\log x}\right)\right)\left|R\left(\frac{x}{[x]}\right)\right| = O(x).$$

Combining these results, we obtain

$$2|R(x)|\log x \le 2\sum_{n \le x}\left|R\left(\frac{x}{n}\right)\right| + O(x\log\log x).$$

Dividing this inequality by $2\log x$ completes the proof of Theorem 9.4. □

Lemma 9.2 *Let $0 < \delta < 1$. There exist numbers $c_0 \ge 1$ and $x_1(\delta) \ge 4$ such that if $x \ge x_1(\delta)$, then there exists an integer n such that*

$$x < n \le e^{c_0/\delta}x$$

and

$$|R(n)| < \delta n.$$

The constant c_0 does not depend on δ.

Proof. By Theorem 6.9,

$$\sum_{n \le x} \frac{1}{n} = \log x + \gamma + r(x) = \log x + O(1),$$

where $|r(x)| < 1/x$. If $1 \le x < x'$, then

$$\sum_{x < n \le x'} \frac{1}{n} = \log\frac{x'}{x} + r'(x),$$

where $|r'(x)| < 2/x$. By Theorem 8.6,

$$\sum_{n \le x} \frac{\vartheta(n)}{n^2} = \log x + O(1).$$

Then

$$\begin{aligned}\sum_{n\leq x}\frac{R(n)}{n^2} &= \sum_{n\leq x}\frac{\vartheta(n)-n}{n^2}\\ &= (\log x+O(1))-(\log x+O(1))\\ &= O(1),\end{aligned}$$

and so

$$\left|\sum_{x<n\leq x'}\frac{R(n)}{n^2}\right| = O(1)$$

for all $1\leq x<x'$. Choose $c_0\geq 1$ such that

$$\left|\sum_{x<n\leq x'}\frac{R(n)}{n^2}\right| < \frac{c_0}{2} \tag{9.8}$$

for all $1\leq x<x'$.

Let $0\delta<1$ and $\rho=e^{c_0/\delta}$. Then $\rho x>ex$. Choose $x_1(\delta)\geq 4$ such that $\log x<\delta x$ for all $x\geq x_1(\delta)$. We must prove that if $x\geq x_1(\delta)$, then there exists an integer $n\in(x,\rho x]$ with $|R(n)|<\delta n$. There are two cases.

In the first case, we assume that either $R(n)\geq 0$ for all integers $n\in(x,\rho x]$, or $R(n)\leq 0$ for all integers $n\in(x,\rho x]$. Then

$$\left|\sum_{x<n\leq\rho x}\frac{R(n)}{n^2}\right| = \sum_{x<n\leq\rho x}\frac{|R(n)|}{n^2} = \sum_{x<n\leq\rho x}\left(\frac{|R(n)|}{n}\right)\frac{1}{n}.$$

If

$$m^*=\min\left\{\frac{|R(n)|}{n}:n\in(x,\rho x]\right\},$$

then

$$\begin{aligned}\frac{c_0}{2} &> \sum_{x<n\leq\rho x}\left(\frac{|R(n)|}{n}\right)\frac{1}{n}\\ &\geq m^*\sum_{x<n\leq\rho x}\frac{1}{n}\\ &> m^*\left(\log\frac{\rho x}{x}-\frac{2}{x}\right)\\ &\geq m^*\left(\frac{c_0}{\delta}-\frac{1}{2}\right)\\ &\geq \frac{c_0 m^*}{2\delta},\end{aligned}$$

and so

$$0 \leq m^* < \delta.$$

There exists an integer $n \in (x, \rho x]$ with $|R(n)|/n = m^*$, and so

$$|R(n)| < \delta n.$$

In the second case, there exist integers $n-1$ and n in the interval $(x, \rho x]$ such that $R(n-1) \neq R(n)$ and $R(n-1)R(n) \leq 0$. Moreover, $n-1 > x \geq x_1(\delta) \geq 4$, and so $n \geq 6$. For every integer $n \geq 2$ we have

$$\begin{aligned} R(n) - R(n-1) &= \vartheta(n) - \vartheta(n-1) - 1 \\ &= \begin{cases} \log n - 1 & \text{if } n \text{ is prime,} \\ -1 & \text{if } n \text{ is composite.} \end{cases} \end{aligned}$$

It follows that if $R(n) < R(n-1)$, then $R(n) - R(n-1) = -1$. Since $R(n) \leq 0 \leq R(n-1)$, we have $|R(n)| \leq 1 < \log n \leq \delta n$. If $R(n-1) < R(n)$, then $R(n-1) \leq 0 \leq R(n)$ and

$$0 \leq R(n) \leq R(n) - R(n-1) = \log n - 1 < \log n < \delta n.$$

In all cases, there exists an integer $n \in (x, \rho x]$ such that $|R(n)| < \delta n$. This completes the proof. □

Lemma 9.3 *Let $c_0 \geq 1$ be the number constructed in Lemma 9.2. and let $0 < \delta < 1$. There exists a number $x_2(\delta)$ such that if $x \geq x_2(\delta)$, then the interval $(x, e^{c_0/\delta}x]$ contains a subinterval $(y, e^{\delta/2}y]$ such that*

$$|R(t)| < 4\delta t$$

for all $t \in (y, e^{\delta/2}y]$.

Proof. We begin with Selberg's formula in the form (9.4). For $x \geq 1$,

$$\sum_{p \leq x} \log p + \sum_{pq \leq x} \frac{\log p \log q}{\log pq} = 2x + O\left(\frac{x}{1 + \log x}\right).$$

For $1 < u \leq t$ we have

$$\begin{aligned} 0 &\leq \sum_{u < p \leq t} \log p \\ &\leq \sum_{u < p \leq t} \log p + \sum_{u < pq \leq t} \frac{\log p \log q}{\log pq} \\ &= 2(t-u) + O\left(\frac{t}{1 + \log t}\right) + O\left(\frac{u}{1 + \log u}\right) \\ &= 2(t-u) + O\left(\frac{t}{1 + \log t}\right), \end{aligned}$$

since the function $t/(1+\log t)$ is increasing for $t \geq 1$. Moreover,

$$\sum_{u<p\leq t} \log p = \vartheta(t) - \vartheta(u) = t - u + R(t) - R(u),$$

and so

$$-(t-u) \leq R(t) - R(u) \leq t - u + O\left(\frac{t}{1+\log t}\right).$$

It follows that if $1 < u \leq t$, then

$$|R(t) - R(u)| \leq t - u + O\left(\frac{t}{1+\log t}\right) \leq t - u + O\left(\frac{t}{\log t}\right).$$

If $1 < t \leq u \leq 2t$, then

$$\begin{aligned} |R(t) - R(u)| &\leq u - t + O\left(\frac{u}{1+\log u}\right) \\ &\leq |t-u| + O\left(\frac{2t}{1+\log 2t}\right) \\ &\leq |t-u| + O\left(\frac{t}{\log t}\right). \end{aligned}$$

In particular, if $u > 4$ and $t/2 \leq u \leq 2t$, then

$$|R(t)| \leq |R(u)| + |t-u| + O\left(\frac{t}{\log t}\right). \tag{9.9}$$

By Lemma 9.2, there is a number $c_0 \geq 1$ such that if $0 < \delta < 1$ and $x \geq x_1(\delta) \geq 4$, then there exists an integer

$$n \in \left(x, e^{c_0/\delta}x\right]$$

with

$$|R(n)| < \delta n.$$

If t is a real number in the interval $[n/2, 2n]$, then $t/2 \leq n \leq 2t$. Since $n > x \geq 4$, we have

$$\log t \geq \log(n/2) > \log(x/2) \geq (\log x)/2,$$

and

$$\begin{aligned} |R(t)| &\leq |R(n)| + |t-n| + O\left(\frac{t}{\log t}\right) \\ &< \delta n + |t-n| + O\left(\frac{t}{\log x}\right) \\ &= t\left(\frac{\delta n}{t} + \left|\frac{t}{n} - 1\right| + O\left(\frac{1}{\log x}\right)\right) \\ &\leq t\left(2\delta + \left|\frac{t}{n} - 1\right| + \frac{c_2}{\log x}\right) \end{aligned}$$

for some constant $c_2 > 0$. If $x \geq x_2(\delta) = \max\left(x_1(\delta), e^{c_2/\delta}\right)$, then

$$|R(t)| < t\left(3\delta + \left|\frac{t}{n} - 1\right|\right).$$

Choose t in the interval

$$e^{-\delta/2}n \leq t \leq e^{\delta/2}n.$$

Then $t \in (n/2, 2n)$ since $e^{\delta/2} < e^{1/2} < 2$. If $t/n \geq 1$, then

$$\left|\frac{t}{n} - 1\right| = \frac{t}{n} - 1 \leq e^{\delta/2} - 1 < \delta,$$

since $e^{\delta/2} < 1 + \delta$ for $0 < \delta < 1$. If $t/n < 1$, then

$$\left|\frac{t}{n} - 1\right| = 1 - \frac{t}{n} \leq 1 - e^{-\delta/2} < e^{\delta/2} - 1 < \delta.$$

Therefore,

$$|R(t)| < 4\delta t.$$

We define the number y as follows. If $e^{\delta/2}n \leq e^{c_0/\delta}x$, let $y = n$. If $e^{\delta/2}n > e^{c_0/\delta}x$, let $y = e^{-\delta/2}n$. Then

$$y = e^{-\delta/2}n > e^{-\delta}e^{c_0/\delta}x = e^{(c_0/\delta)-\delta}x > x,$$

since $c_0/\delta > c_0 \geq 1 > \delta$. In both cases,

$$(y, e^{\delta/2}y] \subseteq (x, e^{c_0/\delta}x]$$

and $|R(t)| < 4\delta t$ for all $t \in (y, e^{\delta/2}y]$. This completes the proof. □

Theorem 9.5 (Prime number theorem) *For Chebyshev's function* $\vartheta(x)$,

$$\vartheta(x) \sim x$$

as $x \to \infty$.

Proof. By Theorem 8.1,

$$\limsup_{x\to\infty} \frac{R(x)}{x} = \limsup_{x\to\infty} \frac{\vartheta(x)}{x} - 1 \leq \log 4 - 1 < 0.4.$$

By Theorem 8.2,

$$\liminf_{x\to\infty} \frac{R(x)}{x} = \liminf_{x\to\infty} \frac{\vartheta(x)}{x} - 1 \geq \log 2 - 1 > -0.4.$$

It follows that there exist numbers M and u_1 such that

$$|R(x)| < Mx \qquad \text{for all } x \geq 1,$$

and

$$|R(x)| < \delta_1 x \quad \text{for all } x \geq u_1,$$

where

$$\delta_1 = 0.4.$$

We shall construct sequences of positive real numbers $\{\delta_m\}_{m=1}^{\infty}$ and $\{\varepsilon_m\}_{m=1}^{\infty}$, such that

$$\delta_1 > \delta_2 > \delta_3 > \cdots$$

and

$$\lim_{m\to\infty} \varepsilon_m = 0. \tag{9.10}$$

Let $m \geq 1$, and suppose that we have constructed the number δ_m. Let $c_0 \geq 1$ be the number defined in Lemma 9.2. Choose ε_m such that

$$0 < \varepsilon_m < 1/m$$

and

$$(1+\varepsilon_m)\left(1 - \frac{\delta_m^2}{256c_0}\right) < 1.$$

We define

$$\delta_{m+1} = (1+\varepsilon_m)\left(1 - \frac{\delta_m^2}{256c_0}\right)\delta_m. \tag{9.11}$$

Then $0 < \delta_{m+1} < \delta_m$. This determines the sequences $\{\delta_m\}_{m=1}^{\infty}$ and $\{\varepsilon_m\}_{m=1}^{\infty}$ inductively.

We shall prove that for every m there exists a number u_m such that

$$|R(x)| < \delta_m x \quad \text{for all } x \geq u_m. \tag{9.12}$$

Let us show that this suffices to prove the prime number theorem. The sequence $\{\delta_m\}_{m=1}^{\infty}$ is a strictly decreasing sequence of positive real numbers, so the sequence converges to some nonnegative number $\delta < 1$. Then (9.10) and (9.11) imply that

$$\delta = \left(1 - \frac{\delta^2}{256c_0}\right)\delta = 0.$$

Inequality (9.12) implies that $R(x) = o(x)$, which is equivalent to the prime number theorem.

We construct the numbers u_m inductively. There exists u_1 such that $|R(x)| < \delta_1 x$ for $x \geq u_1$. Suppose that u_m has been determined. We shall prove that there exists a number u_{m+1} such that $|R(x)| < \delta_{m+1}x$ for all $x \geq u_{m+1}$.

Define

$$\delta'_m = \frac{\delta_m}{8}$$

and

$$\rho = e^{c_0/\delta'_m}.$$

Let $x_2(\delta'_m)$ be the number constructed in Lemma 9.3, and let

$$x_3(m) = \max\left(x_2(\delta'_m), u_m\right).$$

If

$$x \geq x_3(m) \geq x_2(\delta'_m),$$

then by Lemma 9.3, every interval $(x, \rho x]$ contains a subinterval $\left(y, e^{\delta'_m/2}y\right]$ such that

$$|R(t)| < 4\delta'_m t = \frac{\delta_m t}{2}$$

for all $t \in \left(y, e^{\delta'_m/2}y\right]$. Let k be the greatest integer such that $\rho^k \leq x/x_3(m)$. Then

$$k \leq \frac{\log x/x_3(m)}{\log \rho} < k+1,$$

and so

$$\begin{aligned} k &= \frac{\log(x/x_3(m))}{\log \rho} + O(1) \\ &= \frac{\delta'_m \log(x/x_3(m))}{c_0} + O(1) \\ &= \frac{\delta_m \log x}{8c_0} + O(1). \end{aligned}$$

By Theorem 9.4,

$$\begin{aligned} |R(x)| &\leq \frac{1}{\log x}\sum_{n\leq x}\left|R\left(\frac{x}{n}\right)\right| + o(x) \\ &= \frac{1}{\log x}\sum_{n\leq \rho^k}\left|R\left(\frac{x}{n}\right)\right| + \frac{1}{\log x}\sum_{\rho^k<n\leq x}\left|R\left(\frac{x}{n}\right)\right| + o(x) \\ &\leq \frac{1}{\log x}\sum_{n\leq \rho^k}\left|R\left(\frac{x}{n}\right)\right| + \frac{Mx}{\log x}\sum_{\rho^k<n\leq x}\frac{1}{n} + o(x) \\ &\leq \frac{1}{\log x}\sum_{n\leq \rho^k}\left|R\left(\frac{x}{n}\right)\right| + o(x), \end{aligned}$$

since

$$\sum_{\rho^k<n\leq x}\frac{1}{n} \leq \sum_{x/(\rho x_3(m))<n\leq x}\frac{1}{n} = \log(\rho x_3(m)) + O(1/x) = O(1).$$

If $1 \leq n \leq \rho^k$, then

$$\frac{x}{n} \geq \frac{x}{\rho^k} \geq x_3(m) \geq u_m$$

and

$$\left|R\left(\frac{x}{n}\right)\right| < \frac{\delta_m x}{n},$$

by the definition of u_m.

For $j = 1, \ldots, k$, we have

$$\frac{x}{\rho^j} \geq \frac{x}{\rho^k} \geq x_3(m) \geq x_2(\delta'_m),$$

and so each interval $\left(\frac{x}{\rho^j}, \frac{x}{\rho^{j-1}}\right]$ contains a subinterval $I_j = \left(y_j, e^{\delta'_m/2} y_j\right]$ such that

$$|R(t)| < 4\delta'_m t = \frac{\delta_m t}{2} \quad \text{for all } t \in I_j.$$

Therefore,

$$\begin{aligned}
\sum_{n \in (\rho^{j-1}, \rho^j]} \left|R\left(\frac{x}{n}\right)\right| &= \sum_{n \in (\rho^{j-1}, \rho^j] \setminus I_j} \left|R\left(\frac{x}{n}\right)\right| + \sum_{n \in I_j} \left|R\left(\frac{x}{n}\right)\right| \\
&< \delta_m x \sum_{n \in (\rho^{j-1}, \rho^j] \setminus I_j} \frac{1}{n} + \frac{\delta_m x}{2} \sum_{n \in I_j} \frac{1}{n} \\
&= \delta_m x \sum_{n \in (\rho^{j-1}, \rho^j]} \frac{1}{n} - \frac{\delta_m x}{2} \sum_{n \in I_j} \frac{1}{n}.
\end{aligned}$$

Then

$$\begin{aligned}
\sum_{n \leq \rho^k} \left|R\left(\frac{x}{n}\right)\right| &= R(x) + \sum_{j=1}^{k} \sum_{n \in (\rho^{j-1}, \rho^j]} \left|R\left(\frac{x}{n}\right)\right| \\
&\leq \delta_m x + \sum_{j=1}^{k} \left(\delta_m x \sum_{n \in (\rho^{j-1}, \rho^j]} \frac{1}{n} - \frac{\delta_m x}{2} \sum_{n \in I_j} \frac{1}{n}\right) \\
&= \delta_m x \sum_{n \leq \rho^k} \frac{1}{n} - \frac{\delta_m x}{2} \sum_{j=1}^{k} \sum_{n \in I_j} \frac{1}{n}.
\end{aligned}$$

We have

$$\begin{aligned}
\delta_m x \sum_{n \leq \rho^k} \frac{1}{n} &= \delta_m x \left(k \log \rho + O\left(\frac{1}{\rho^k}\right)\right) \\
&= \delta_m x \log x + O(x).
\end{aligned}$$

Moreover,

$$\sum_{n\in I_j}\frac{1}{n} = \sum_{n\in(y_j,e^{\delta'_m/2}y_j]}\frac{1}{n} = \frac{\delta'_m}{2} + O\left(\frac{1}{y_j}\right) = \frac{\delta'_m}{2} + O\left(\frac{\rho^j}{x}\right),$$

and so

$$\begin{aligned}\sum_{j=1}^{k}\sum_{n\in I_j}\frac{1}{n} &= \frac{\delta'_m k}{2} + O\left(\sum_{j=1}^{k}\frac{\rho^j}{x}\right)\\ &= \frac{\delta'_m}{2}\left(\frac{\delta_m\log x}{8c_0} + O(1)\right) + O(1)\\ &= \frac{\delta_m^2\log x}{128c_0} + O(1),\end{aligned}$$

since

$$\sum_{j=1}^{k}\frac{\rho^j}{x} = \frac{\rho(\rho^k-1)}{x(\rho-1)} < \frac{2\rho^k}{x} \leq \frac{2}{x_3(m)} = O(1).$$

Therefore,

$$\frac{\delta_m x}{2}\sum_{j=1}^{k}\sum_{n\in I_j}\frac{1}{n} = \frac{\delta_m^3 x\log x}{256c_0} + O(x).$$

Combining these results, we obtain, for $x \geq x_3(m)$,

$$\begin{aligned}\sum_{n\leq\rho^k}\left|R\left(\frac{x}{n}\right)\right| &\leq (\delta_m x\log x + O(x)) - \left(\frac{\delta_m^3 x\log x}{256c_0} + O(x)\right)\\ &= \left(1-\frac{\delta_m^2}{256c_0}\right)\delta_m x\log x + O(x),\end{aligned}$$

and

$$\begin{aligned}|R(x)| &\leq \frac{1}{\log x}\sum_{n\leq\rho^k}\left|R\left(\frac{x}{n}\right)\right| + o(x)\\ &= \left(1-\frac{\delta_m^2}{256c_0}\right)\delta_m x + o(x).\end{aligned}$$

We choose u_{m+1} sufficiently large that for all $x \geq u_{m+1}$ we have

$$o(x) < \varepsilon_m\left(1-\frac{\delta_m^2}{256c_0}\right)\delta_m x.$$

Then

$$|R(x)| < (1+\varepsilon_m)\left(1-\frac{\delta_m^2}{256c_0}\right)\delta_m x = \delta_{m+1}x.$$

This completes the proof of the prime number theorem. □

Exercises

1. Let p_n denote the nth prime number. Prove that $p_n \sim n \log n$.

2. Prove that
$$\lim_{n\to\infty} \frac{p_{n+1}}{p_n} = 1.$$

3. Let $\delta > 0$. Prove that
$$\vartheta((1+\delta)x) - \vartheta(x) \sim \delta x.$$
This implies that there is a prime between x and $(1+\delta)x$ for all sufficiently large x.

4. Prove that
$$\pi((1+\delta)x) - \pi(x) \sim \frac{\delta x}{\log x}.$$

5. Prove that
$$\pi(2x) - \pi(x) \sim \pi(x).$$

6. Let p_n denote the nth prime number, so that $p_1 = 2$, $p_2 = 3, \ldots$. Prove that the asymptotic formula $p_n \sim n \log n$ implies the prime number theorem.

7. Deduce Selberg's formula (9.3) from the prime number theorem.

8. Let $\delta_1 = 2$. For every $m \geq 1$ define
$$\delta_{m+1} = \delta_m \left(1 - \frac{\delta_m^2}{256c_0}\right).$$
Prove that
$$0 < \delta_m \ll \frac{1}{\sqrt{m}}.$$

9.4 Integers with k Prime Factors

For any positive integer n, the arithmetic functions $\omega(n)$ and $\Omega(n)$ are defined as follows: $\omega(n) = k$ if n is divisible by exactly k different primes, and $\Omega(n) = \ell$ if n is the product of ℓ not necessarily distinct primes. If $n = p_1^{a_1} \cdots p_k^{a_k}$, where $p_1, \ldots, p_k$ are pairwise distinct prime numbers and $a_1, \ldots, a_k$ are positive integers, then $\omega(n) = k$ and $\Omega(n) = a_1 + \cdots + a_k$.

Let $\pi_k(x)$ denote the number of positive integers n not exceeding x that can be written as the product of exactly k distinct primes,
$$\pi_k(x) = \sum_{\substack{n \leq x \\ \omega(n)=\Omega(n)=k}} 1.$$

Let $\pi_k^*(x)$ denote the number of positive integers n not exceeding x that can be written as the product of exactly k not necessarily distinct prime numbers,

$$\pi_k^*(x) = \sum_{\substack{n \leq x \\ \Omega(n)=k}} 1.$$

Our goal is the following asymptotic estimate for the number of integers with exactly k prime divisors:

$$\pi_k(x) \sim \pi_k^*(x) \sim \frac{x(\log\log x)^{k-1}}{(k-1)!\log x}.$$

This is a generalization of the prime number theorem, since $\pi_1(x) = \pi_1^*(x) = \pi(x) \sim x/\log x$.

Let $\mathbf{P} = \{2, 3, 5, \ldots\}$ be the set of prime numbers, and let $\mathbf{P}^k$ be the set of all ordered k-tuples of primes. Let $r_k(n)$ denote the number of representations of n as an ordered product of k primes, that is,

$$r_k(n) = \sum_{\substack{(p_1,\ldots,p_k)\in\mathbf{P}^k \\ p_1\cdots p_k=n}} 1.$$

Since every positive integer is uniquely (up to order) a product of primes, we have

$$0 \leq r_k(n) \leq k! \qquad \text{for all } n \geq 1.$$

Moreover, $r_k(n) = k!$ if and only if $\omega(n) = \Omega(n) = k$, and $0 < r_k(n) < k!$ if and only if $\omega(n) < \Omega(n) = k$.

Theorem 9.6 *For $k \geq 1$, let*

$$\Pi_k^*(x) = \sum_{n\leq x} r_k(n) = \sum_{\substack{(p_1,\ldots,p_k)\in\mathbf{P}^k \\ p_1\cdots p_k\leq x}} 1.$$

Then

$$k!\pi_k(x) \leq \Pi_k^*(x) \leq k!\pi_k^*(x) \ll x.$$

For $k \geq 2$,

$$0 \leq \pi_k^*(x) - \pi_k(x) \leq \Pi_{k-1}^*(x).$$

Proof. We have

$$\Pi_k^*(x) = \sum_{n\leq x} r_k(n) \leq k! \sum_{\substack{n\leq x \\ r_k(n)>0}} 1 = k!\pi_k^*(x) \leq k!x \ll x$$

and

$$\Pi_k^*(x) = \sum_{n\leq x} r_k(n) \geq k! \sum_{\substack{n\leq x \\ r_k(n)=k!}} 1 = k!\pi_k(x).$$

Let $k \geq 2$. The function $\pi_k^*(x) - \pi_k(x)$ counts the number of positive integers $n \leq x$ that can be written as a product of k primes but not as a product of k distinct primes. Every such integer is of the form $n = p_1 \cdots p_{k-2} p_{k-1}^2$. Therefore,

$$\begin{aligned} \pi_k^*(x) - \pi_k(x) &\leq \sum_{\substack{(p_1,\ldots,p_{k-1}) \in \mathbf{P}^{k-1} \\ p_1 \cdots p_{k-1}^2 \leq x}} 1 \\ &\leq \sum_{\substack{(p_1,\ldots,p_{k-1}) \in \mathbf{P}^{k-1} \\ p_1 \cdots p_{k-1} \leq x}} 1 \\ &= \Pi_{k-1}^*(x). \end{aligned}$$

This completes the proof. □

Theorem 9.7 *Let $S_0(x) = 1$. For $k \geq 1$, let*

$$S_k(x) = \sum_{\substack{(p_1,\ldots,p_k) \in \mathbf{P}^k \\ p_1 \cdots p_k \leq x}} \frac{1}{p_1 \cdots p_k} = \sum_{n \leq x} \frac{r_k(n)}{n}.$$

Then

$$S_k(x) \sim (\log\log x)^k$$

and

$$S_k(x) = \sum_{p \leq x} \frac{1}{p} S_{k-1}\left(\frac{x}{p}\right).$$

Proof. By Theorem 8.7,

$$S_1(x) = \sum_{p \leq x} \frac{1}{p} \sim \log\log x$$

and so

$$S_1(x^{1/k}) \sim \log\log x^{1/k} \sim \log\log x$$

for all $k \geq 1$. Since

$$\begin{aligned} \left(S_1\left(x^{1/k}\right)\right)^k &= \left(\sum_{p \leq x^{1/k}} \frac{1}{p}\right)^k = \sum_{\substack{(p_1,\ldots,p_k) \in \mathbf{P}^k \\ p_i \leq x^{1/k}}} \frac{1}{p_1 \cdots p_k} \\ &\leq \sum_{\substack{(p_1,\ldots,p_k) \in \mathbf{P}^k \\ p_1 \cdots p_k \leq x}} \frac{1}{p_1 \cdots p_k} = S_k(x) \\ &\leq \left(\sum_{p \leq x} \frac{1}{p}\right)^k = S_1(x)^k, \end{aligned}$$

it follows that

$$S_k(x) \sim (\log\log x)^k.$$

Also,

$$\begin{aligned} S_k(x) &= \sum_{\substack{(p_1,\ldots,p_{k-1},p_k)\in\mathbf{P}^k \\ p_1\cdots p_{k-1}p_k\le x}} \frac{1}{p_1\cdots p_{k-1}p_k} \\ &= \sum_{p_k\le x}\frac{1}{p_k} \sum_{\substack{(p_1,\ldots,p_{k-1})\in\mathbf{P}^{k-1} \\ p_1\cdots p_{k-1}\le x/p_k}} \frac{1}{p_1\cdots p_{k-1}} \\ &= \sum_{p_k\le x}\frac{1}{p_k} S_{k-1}\left(\frac{x}{p_k}\right). \end{aligned}$$

This completes the proof. □

Theorem 9.8 *For $k \geq 1$, let*

$$\vartheta_k(x) = \sum_{\substack{(p_1,\ldots,p_k)\in\mathbf{P}^k \\ p_1\cdots p_k\le x}} \log p_1\cdots p_k.$$

Then

$$\vartheta_k(x) \sim kx(\log\log x)^{k-1}.$$

Proof. For $j = 1, \ldots, k+1$, let

$$p_1\cdots\hat{p_j}\cdots p_{k+1} = \prod_{\substack{i=1 \\ i\neq j}}^{k+1} p_i.$$

Then

$$\sum_{j=1}^{k+1}\log p_1\cdots\hat{p_j}\cdots p_{k+1} = \log(p_1\cdots p_{k+1})^k = k\log p_1\cdots p_{k+1},$$

and so, by Exercise 4,

$$\begin{aligned} k\vartheta_{k+1}(x) &= \sum_{\substack{(p_1,\ldots,p_{k+1})\in\mathbf{P}^{k+1} \\ p_1\cdots p_{k+1}\le x}} k\log p_1\cdots p_{k+1} \\ &= \sum_{\substack{(p_1,\ldots,p_{k+1})\in\mathbf{P}^{k+1} \\ p_1\cdots p_{k+1}\le x}} \sum_{j=1}^{k+1}\log p_1\cdots\hat{p_j}\cdots p_{k+1} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\substack{(p_1,\ldots,p_{k+1})\in\mathbf{P}^{k+1}\\ p_1\cdots p_{k+1}\leq x}} (k+1)\log p_1\cdots p_k \\
&= (k+1)\sum_{p_{k+1}\leq x}\ \sum_{\substack{(p_1,\ldots,p_k)\in\mathbf{P}^{k}\\ p_1\cdots p_k\leq x/p_{k+1}}} \log p_1\cdots p_k \\
&= (k+1)\sum_{p\leq x}\vartheta_k\left(\frac{x}{p}\right).
\end{aligned}
$$

For $k \geq 1$, let

$$F_k(x) = \vartheta_k(x) - kxS_{k-1}(x).$$

Then

$$
\begin{aligned}
kF_{k+1}(x) &= k\vartheta_{k+1}(x) - k(k+1)xS_k(x) \\
&= (k+1)\sum_{p\leq x}\vartheta_k\left(\frac{x}{p}\right) - k(k+1)\sum_{p\leq x}\frac{x}{p}S_{k-1}\left(\frac{x}{p}\right) \\
&= (k+1)\sum_{p\leq x}\left(\vartheta_k\left(\frac{x}{p}\right) - \frac{kx}{p}S_{k-1}\left(\frac{x}{p}\right)\right) \\
&= (k+1)\sum_{p\leq x}F_k\left(\frac{x}{p}\right).
\end{aligned}
$$

We shall prove by induction that

$$F_k(x) = o\left(x(\log\log x)^{k-1}\right). \tag{9.13}$$

For $k = 1$,

$$F_1(x) = \vartheta(x) - x = o(x)$$

is the prime number theorem. Suppose that (9.13) is true for some $k \geq 1$. Let $\varepsilon > 0$. There exists $x_0(\varepsilon)$ such that

$$|F_k(x)| \leq \varepsilon x(\log\log x)^{k-1}$$

for all $x \geq x_0 = x_0(\varepsilon)$, and so

$$\sum_{p\leq x/x_0} F_k\left(\frac{x}{p}\right) \leq \varepsilon x(\log\log x)^{k-1}\sum_{p\leq x/x_0}\frac{1}{p} \leq 2\varepsilon x(\log\log x)^k$$

for $x \geq x_1 = x_1(\varepsilon) \geq x_0$. Since the functions $\vartheta_k(x)$ and $S_{k-1}(x)$ are nonnegative and increasing for $x \geq 1$, it follows that $F_k(x)$ is bounded on any finite interval, and so there exists a constant $M_1 = M_1(\varepsilon)$ such that

$$|F_k(x)| \leq M_1 \qquad \text{for } 1 \leq x \leq x_1.$$

Therefore,

$$\begin{aligned} kF_{k+1}(x) &= (k+1)\sum_{p\leq x} F_k\left(\frac{x}{p}\right) \\ &= (k+1)\sum_{p\leq x/x_0} F_k\left(\frac{x}{p}\right) + (k+1)\sum_{x/x_0<p\leq x} F_k\left(\frac{x}{p}\right) \\ &\leq 2(k+1)\varepsilon x(\log\log x)^k + (k+1)M_1\pi(x) \\ &\leq 4k\varepsilon x(\log\log x)^k + 2kM_1 x. \end{aligned}$$

Dividing by k, we obtain

$$F_{k+1}(x) \ll \varepsilon x(\log\log x)^k$$

for all sufficiently large x. This proves (9.13). It follows that

$$\begin{aligned} \vartheta_k(x) &= kxS_{k-1}(x) + F_k(x) \\ &= kx(\log\log x)^{k-1} + o\left(x(\log\log x)^{k-1}\right) \\ &\sim kx(\log\log x)^{k-1}. \end{aligned}$$

This completes the proof. □

Theorem 9.9 *For $k \geq 1$,*

$$\pi_k(x) \sim \pi_k^*(x) \sim \frac{x(\log\log x)^{k-1}}{k\log x}.$$

Proof. This follows from Theorem 9.8 by partial summation. We have

$$\vartheta_k(x) = \sum_{\substack{(p_1,\ldots,p_k)\in\mathbf{P}^k \\ p_1\cdots p_k\leq x}} \log p_1\cdots p_k = \sum_{n\leq x} r_k(n)\log n,$$

and, by Theorem 9.6, the arithmetic function $r_k(n)$ has mean value

$$\Pi_k^*(x) = \sum_{n\leq x} r_k(n) = O(x).$$

Then

$$\begin{aligned} \vartheta_k(x) &= \Pi_k^*(x)\log x - \int_1^x \frac{\Pi_k^*(t)dt}{t} \\ &= \Pi_k^*(x)\log x + O(x). \end{aligned}$$

It follows that

$$\Pi_k^*(x) = \frac{\vartheta_k(x)}{\log x} + O\left(\frac{x}{\log x}\right) \sim \frac{kx(\log\log x)^{k-1}}{\log x}.$$

For $k \geq 2$,

$$\Pi_{k-1}^*(x) = o\left(\Pi_k^*(x)\right).$$

By Theorem 9.6,

$$\Pi_k^*(x) \leq k!\pi_k^*(x) \leq k!\pi_k(x) + k!\Pi_{k-1}^*(x) \leq \Pi_k^*(x) + k!\Pi_{k-1}^*(x),$$

and so

$$\pi_k^*(x) \sim \pi_k(x) \sim \frac{\Pi_k^*(x)}{k!} \sim \frac{x(\log\log x)^{k-1}}{(k-1)!\log x}.$$

This completes the proof. □

Exercises

1. For every positive integer n, let $r_k(n)$ denote the number of k-tuples of prime numbers $(p_1, \ldots, p_k)$ such that $n = p_1 \cdots p_k$. Compute $r_3(n)$ for $n \leq 50$.

2. Compute $r_4(n)$ for $n \leq 100$.

3. Let $\sigma > 1$. Prove that

$$\sum_{n=1}^{\infty} \frac{r_k(n)}{n^\sigma} = \left(\sum_{p\in\mathbf{P}} \frac{1}{p^\sigma}\right)^k.$$

4. Prove that

$$\sum_{\substack{(p_1,\ldots,p_{k+1})\in\mathbf{P}^{k+1} \\ p_1\cdots p_{k+1}\leq x}} \sum_{j=1}^{k+1} \log p_1 \cdots \hat{p_j} \cdots p_{k+1} = \sum_{\substack{(p_1,\ldots,p_{k+1})\in\mathbf{P}^{k+1} \\ p_1\cdots p_{k+1}\leq x}} (k+1)\log p_1 \cdots p_k.$$

5. Let x_k be the smallest number such that $\pi_k(x_k) > 0$. Prove that for every $\varepsilon > 0$ there exists an integer $k_0 = k_0(\varepsilon)$ such that if $k \geq k_0$, then

$$k^{(1-\varepsilon)k} < x_k < k^{(1+\varepsilon)k}.$$

9.5 Notes

In a lecture delivered to the Mathematical Society of Copenhagen in 1921, Hardy said,

> No elementary proof of the prime number theorem is known, and one may ask whether it is reasonable to expect one. Now we know that the theorem is roughly equivalent to a theorem about an analytic function, the theorem that Riemann's zeta function has no roots on a certain line. A proof of such a theorem, not fundamentally dependent upon the ideas of the theory of functions, seems to me extraordinarily unlikely. It is rash to assert that a mathematical theorem *cannot* be proved in a particular way; but one thing seems quite clear. We have certain views about the logic of the theory; we think that some theorems, as we say "lie deep" and others nearer to the surface. If anyone produces an elementary proof of the prime number theorem, he will show that these views are wrong, that the subject does not hang together in the way we have supposed, and that it is time for the books to be cast aside and for the theory to be rewritten.
>
> G. H. Hardy, quoted in Bohr [11]

In 1949, in a review of the Erdős and Selberg elementary proofs of the prime number theorem, Ingham wrote,

> What Selberg and Erdős do is to deduce the PNT directly ... without the explicit intervention of the analytical fact .... How far the practical effects of this revolution of ideas will penetrate into the structure of the subject, and how much of the theory will ultimately have to be rewritten, it is too early to say.
>
> A. E. Ingham [71]

The prime number theorem was proved independently in 1896 by J. Hadamard [46] and C.-J. de la Vallée Poussin[23]. Their proofs applied complex function theory to the Riemann zeta function. Ingham's classic monograph, *The Distribution of Prime Numbers* [70], published in 1932, contains an analytic proof of the prime number theorem.

The elementary proof of the prime number theorem was discovered in 1948 at the Institute for Advanced Study in Princeton. In March 1948, Selberg discovered his famous formula (Theorems 9.2 and 9.3) and gave an elementary proof of Dirichlet's theorem on primes in arithmetic progressions (Theorem 10.9). By April 1948, he knew that $A + a = 2$ (Exercises 4

and 5 in Section 9.2), and that the prime number theorem is equivalent to $A = a = 1$. In a letter to H. Weyl that is dated September 16, 1948, Selberg[1] wrote:

> In May I wrote down a sketch to the paper on Dirichlet's theorem, during June I did nothing except preparations to the trip to Canada. Then around the beginning of July, Turán asked me if I could give him my notes on the Dirichlet theorem so he could see it, he was going away soon, and probably would have left when I returned from Canada. I not only agreed to do this, but as I felt very much attached to Turán I spent some days going through the proof with him. In this connection I mentioned the *fundamental theorem* to him.... However, I did not tell him the proof of the formula, nor about the consequences it might have and my ideas in this connection.... I then left for Canada and returned after 9 days just as Turán was leaving. It turned out that Turán had given a seminar on my proof of the Dirichlet theorem where Erdős, Chowla, and Straus had been present, I had of course no objection to this, since it concerned something that was already finished from my side, though it was not published. In connection with this Turán had also mentioned, at least to Erdős, the *fundamental formula*....

In a letter to D. Goldfeld that is dated January 6, 1988, Selberg wrote:

> July 14, 1948 was a Wednesday, and on Thursday, July 15 I met Erdős and heard that he was trying to prove $p_{n+1}/p_n \to 1$.... Friday evening or it may have been Saturday morning, Erdős had his proof ready and told me about it. Sunday afternoon (July 18) I used his result (which was stronger than $p_{n+1}/p_n \to 1$, he had proved that between x and $x(1+\delta)$there are more that $c(\delta)x/\log x$ primes for $x > x_0(\delta)$, the weaker result would not have been sufficient for me) to get my first proof of PNT. I told Erdős about it the next morning (Monday, July 19). He then suggested that we should talk about it that evening in the seminar room in Fuld Hall....

Erdős records the history of the first elementary proof of the prime number theorem in the same way:

> In the course of several important researches in elementary number theory A. Selberg proved some months ago the follow-

[1] This and the following extract from Selberg's correspondence appear in Goldfeld's historical note [38]

> ing important asymptotic formula:
>
> $$\sum_{p\leq x}(\log p)^2 + \sum_{pq\leq x}\log p\log q = 2x\log x + O(x), \qquad (9.14)$$
>
> where p and q run over the primes....
>
> Using (9.14) I proved that $p_{n+1}/p_n \to 1$ as $n \to \infty$. In fact I proved the following slightly stronger result: To every c there exists a positive $\delta(c)$, so that for x sufficiently large we have
>
> $$\pi[x(1+c)] - \pi(x) > \delta(c)x/\log x \qquad (9.15)$$
>
> where $\pi(x)$ is the number of primes not exceeding x.
>
> I communicated this proof of (9.15) to Selberg, who, two days later, using (9.14), (9.15) and the ideas of the proof of (9.15), deduced the prime number theorem....
>
> Erdős [34, pp. 374–375]

Both Erdős [34] and Selberg [128] subsequently gave independent elementary proofs of the prime number theorem. These proofs all use Selberg's original formula, as well as ideas that Erdős introduced in his proof of (9.15).

Number theorists of Hardy's and Ingham's generation believed that there could be no elementary proof of the prime number theorem. They were also convinced that if, by some miracle, an elementary proof were discovered, then the ideas in that proof would lead to tremendous progress in our knowledge of the distribution of prime numbers and the zeros of the zeta function. Both statements are false. Erdős and Selberg produced elementary proofs, but their beautiful method has not led to any extraordinary new discoveries in number theory or analysis.

> The elementary proof has so far not produced the exciting innovations in number theory that many of us expected to follow. So, what we witnessed in 1948, may in the course of time prove to have been a brilliant but somewhat incidental achievement without the historic significance it then appeared to have. My own inclination is to believe that it was the beginning of important new ideas not yet fully understood and that its importance will grow over the years.
>
> E. G. Straus [136]

The elementary proof of the prime number theorem that appears in this chapter is the proof in Selberg's original paper [128]. Postnikov and Romanov [115, 116] give a similar elementary proof in terms of the Möbius

function. Daboussi [18] and Hildebrand [67] obtained elementary proofs of the prime number theorem that do not depend on Selberg's formula. Diamond [24] has written a careful survey of elementary methods in prime number theory.

For more recent developments in prime number theory, see Tenenbaum and Mendès-France, *The Prime Numbers and Their Distribution* [140]. D. J. Newman has recently published a simple analytic proof (Newman [112], Zagier [159]).

The asymptotic formula for the number of integers with exactly k prime factors is based on work of E. M. Wright (see Hardy and Wright [60, pp. 368–370]).

The most important unsolved problem in mathematics is the *Riemann hypothesis*. It can be expressed in terms of the distribution of prime numbers. By Exercise 2 in Section 9.2, the logarithmic integral $\mathrm{li}(x)$ is asymptotic to $x/\log x$, and so the prime number theorem can be restated in the form

$$\pi(x) \sim \mathrm{li}(x).$$

The Riemann hypothesis is an assertion about the size of the error term in the prime number theorem, namely, that

$$\pi(x) = \mathrm{li}(x) + O\left(x^{1/2+\varepsilon}\right)$$

for every $\varepsilon > 0$.

10

Primes in Arithmetic Progressions

10.1 Dirichlet Characters

Dirichlet's theorem states that if $m \geq 1$ and a are relatively prime integers, then the arithmetic progression $mk + a$ contains infinitely many primes, that is, there exist infinitely many primes p of the form $p = mk + a$. Equivalently, the congruence class $a \pmod{m}$ contains infinitely many prime numbers. For example, there are infinitely many primes p such that $p \equiv 3 \pmod{4}$, and there are infinitely many primes p such that $p \equiv 5 \pmod{6}$, by Exercises 8 and 9 in Section 1.5.

We begin by constructing a class of periodic functions called *Dirichlet characters* whose domain is the set of integers.

Let m be a positive integer and let $\mathbf{Z}/m\mathbf{Z}$ be the ring of congruence classes modulo m. The additive group of this ring is cyclic of order m, and its dual group is also cyclic of order m. A character of the additive group $\mathbf{Z}/m\mathbf{Z}$ is called an *additive character modulo m.*

Let ζ be a primitive mth root of unity. If ψ is an additive character modulo m, then there exists a unique integer $a \in \{0, 1, 2, \ldots, m-1\}$ such that

$$\psi(k + m\mathbf{Z}) = \zeta^{ak}.$$

Choosing the primitive mth root of unity $\zeta = \exp(2\pi i/m)$, we have

$$\psi_a(k + m\mathbf{Z}) = \exp\left(\frac{2\pi iak}{m}\right) = e_m(ak).$$

Associated to the additive character ψ_a is a complex-valued function ψ'_a on the integers that is defined by

$$\psi'_a(k) = \psi_a(k + m\mathbf{Z}).$$

We let ψ_a denote both the additive character modulo m and its associated function on the integers.

The group of units in the ring of integers modulo m is the multiplicative group $(\mathbf{Z}/m\mathbf{Z})^\times$ of order $\varphi(m)$, where $\varphi(m)$ is the Euler φ-function. A character of this group is called a *multiplicative character modulo* m. The *principal character* χ_0 modulo m is the multiplicative character defined by $\chi_0(a + m\mathbf{Z}) = 1$ for all $a + m\mathbf{Z} \in (\mathbf{Z}/m\mathbf{Z})^\times$.

For every multiplicative character χ, we have

$$\chi(-1 + m\mathbf{Z})^2 = \chi(1 + m\mathbf{Z}) = 1,$$

and so

$$\chi(-1 + m\mathbf{Z}) = \pm 1.$$

The character χ is called *even* if $\chi(-1+m\mathbf{Z}) = 1$ and *odd* if $\chi(-1+m\mathbf{Z}) = -1$.

A multiplicative character modulo m is called *real* if it is real-valued. Since the only real roots of unity are ± 1, it follows that if χ is a real character, then $\chi(a + m/Z) = \pm 1$ for all $(a, m) = 1$. The character χ is called *complex* if $\chi(a + m\mathbf{Z})$ is not real for some congruence class $a + m\mathbf{Z}$.

Let χ be a multiplicative character modulo m. We extend χ to the nonunits of the ring $\mathbf{Z}/m\mathbf{Z}$ by setting $\chi(a + m\mathbf{Z}) = 0$ whenever $(a, m) = 1$.

For every odd prime p, the Legendre symbol $\left(\frac{\cdot}{p}\right)$ defines a real multiplicative character χ modulo p by

$$\chi(a + p\mathbf{Z}) = \left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p, \\ 0 & \text{if } (a, p) > 1. \end{cases}$$

By Theorem 3.14, this character is even if $p \equiv 1 \pmod 4$ and odd if $p \equiv 3 \pmod 4$.

Corresponding to every multiplicative character χ modulo m there is a complex-valued function χ' on the integers defined by

$$\chi'(a) = \chi(a + m\mathbf{Z}).$$

The function $\chi' : \mathbf{Z} \to \mathbf{C}$ is called a *Dirichlet character modulo* m.

A Dirichlet character χ' modulo m has the following properties:

(i) The function χ' has period m, that is, if $a \equiv b \pmod m$, then $\chi'(a) = \chi'(b)$.

(ii) The support of χ' is the set of integers relatively prime to m, that is, $\chi'(a) \neq 0$ if and only if $(a, m) = 1$.

(iii) χ' is completely multiplicative, that is, $\chi'(ab) = \chi'(a)\chi'(b)$ for all integers a and b.

Conversely, every complex-valued function χ' on the integers that satisfies properties (i), (ii), and (iii) is a Dirichlet character modulo m, and the multiplicative character χ modulo m that corresponds to χ' is defined by

$$\chi(a + m\mathbf{Z}) = \chi'(a).$$

From now on, we shall use χ to denote both a multiplicative character modulo m and the corresponding Dirichlet character modulo m.

The principal Dirichlet character χ_0 modulo m is defined by $\chi_0(a) = 1$ if $(a, m) = 1$ and $\chi_0(a) = 0$ if $(a, m) \geq 2$. In particular, the principal Dirichlet character modulo 1 satisfies $\chi_0(a) = 1$ for all integers a.

A Dirichlet character modulo m is called real, complex, even, or odd precisely when the corresponding multiplicative character modulo m is real, complex, even, or odd, respectively.

We can state the orthogonality relations for Dirichlet characters as follows.

Theorem 10.1 (Orthogonality relations) *Let $\sum_{a \pmod m}$ denote the sum over a complete set of residue classes modulo m, and let $\sum_{\chi \pmod m}$ denote the sum over the $\varphi(m)$ Dirichlet characters modulo m. If χ is a Dirichlet character modulo m, then*

$$\sum_{a \pmod m} \chi(a) = \begin{cases} \varphi(m) & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0. \end{cases}$$

If a is an integer, then

$$\sum_{\chi \pmod m} \chi(a) = \begin{cases} \varphi(m) & \text{if } a \equiv 1 \pmod m, \\ 0 & \text{if } a \not\equiv 1 \pmod m. \end{cases}$$

Proof. This is simply Theorem 4.6 applied to the multiplicative group $(\mathbf{Z}/m\mathbf{Z})^{\times}$. □

Theorem 10.2 (Orthogonality relations) *Let $\sum_{a \pmod m}$ denote the sum over a complete set of residue classes modulo m, and let $\sum_{\chi \pmod m}$ denote the sum over the $\varphi(m)$ Dirichlet characters modulo m. If χ_1 and χ_2 are Dirichlet characters modulo m, then*

$$\sum_{a \pmod m} \chi_1(a)\overline{\chi_2}(a) = \begin{cases} \varphi(m) & \text{if } \chi_1 = \chi_2, \\ 0 & \text{if } \chi_1 \neq \chi_2. \end{cases}$$

If a and b are integers, then

$$\sum_{\chi \pmod m} \chi(a)\overline{\chi}(b) = \begin{cases} \varphi(m) & \text{if } (a,m) = (b,m) = 1 \text{ and } a \equiv b \pmod m, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. This is Theorem 4.7 applied to the multiplicative group $(\mathbf{Z}/m\mathbf{Z})^\times$. □

Let d and m be positive integers such that d divides m. There is a natural ring homomorphism

$$\pi : \mathbf{Z}/m\mathbf{Z} \to \mathbf{Z}/d\mathbf{Z}$$

defined by

$$\pi(a + m\mathbf{Z}) = a + d\mathbf{Z}.$$

If $(a,m) = 1$, then $(a,d) = 1$ and so π induces a group homomorphism $\pi : (\mathbf{Z}/m\mathbf{Z})^\times \to (\mathbf{Z}/d\mathbf{Z})^\times$ on the unit groups of these rings. Let λ be a multiplicative character modulo d. The composition of the maps

$$(\mathbf{Z}/m\mathbf{Z})^\times \xrightarrow{\pi} (\mathbf{Z}/d\mathbf{Z})^\times \xrightarrow{\lambda} \mathbf{C}^\times$$

induces a multiplicative character χ modulo m defined by

$$\chi = \lambda\pi,$$

and so

$$\chi(a + m\mathbf{Z}) = \lambda(a + d\mathbf{Z}).$$

This character is called an *induced character*. A character χ modulo m is called a *primitive character* if it is not induced from a character modulo d for any proper divisor d of m.

Alternatively, we can define induced characters by means of Dirichlet characters modulo m. Let d and m be positive integers such that d divides m. If λ is a Dirichlet character modulo d, then we can define a Dirichlet character χ modulo m by the formula

$$\chi(a) = \begin{cases} \lambda(a) & \text{if } (a,m) = 1, \\ 0 & \text{if } (a,m) \neq 1. \end{cases}$$

Let d, k, and m be positive integers such that d divides k and k divides m, and let λ, σ, and χ be Dirichlet (or multiplicative) characters modulo d, k, and m, respectively. If the character λ induces σ and the character σ induces χ, then λ induces χ.

There is a unique Dirichlet character modulo 1; it is the constant function $\lambda(a) = 1$ for all integers a. For every $m \geq 2$, the character λ induces the principal character χ_0 modulo m.

Exercises

1. Construct all of the Dirichlet characters modulo 5.
2. Prove that the nontrivial Dirichlet character modulo 6 is induced by a primitive Dirichlet character modulo 3.
3. Construct all Dirichlet characters modulo 4 and modulo 8. Find the primitive characters.
4. Let m and d be positive integers such that d divides m. Let λ be a Dirichlet character modulo d, and let χ be the Dirichlet character modulo m induced by λ. Prove that $\chi(a) = \lambda(a)\chi_0(a)$, where χ_0 is the principal character modulo m.
5. Let χ be the principal Dirichlet character modulo m. Prove that
$$\sum_{n=a}^{b} \chi(n) \geq \left[\frac{b-a+1}{m}\right] \varphi(m)$$
for all integers a and b.
6. Let χ be a nonprincipal Dirichlet character modulo m. Prove that
$$\sum_{n=a}^{b} \chi(n) < \varphi(m)$$
for all integers a and b.
7. Prove that for every integer a,
$$\sum_{\chi} \chi(a) = \begin{cases} \varphi(m) & \text{if } a \equiv 1 \pmod{m}, \\ 0 & \text{if } a \not\equiv 1 \pmod{m}, \end{cases}$$
where the summation runs over all of the Dirichlet characters modulo m.
8. Let $\varphi^*(m)$ denote the number of primitive characters modulo m. Prove that
$$\varphi(m) = \sum_{d \mid m} \varphi^*(d),$$
where $\varphi(m)$ is the Euler phi function.
9. Prove that $\varphi^*(m)$ is a multiplicative function and that
$$\varphi^*(m) = \sum_{d \mid m} \mu\left(\frac{m}{d}\right) \varphi(d).$$
10. Prove that
$$\varphi^*(m) = m \prod_{p \| m} \left(1 - \frac{2}{p}\right) \prod_{p^2 \mid m} \left(1 - \frac{1}{p}\right)^2.$$

10.2 Dirichlet L-Functions

We begin by introducing a class of functions that are analytic on half-planes of the complex plane. The proof of Dirichlet's theorem, however, involves only routine partial summations of the infinite series and infinite product representations of these functions on the positive real axis. We do not use complex function theory, and, indeed, it would suffice to consider the L-functions only for $\sigma > 0$.

Let χ be a Dirichlet character modulo m. The *Dirichlet L-function* associated with χ is the function

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where

$$s = \sigma + it$$

is a complex number with real part $\Re(s) = \sigma$ and imaginary part $\Im(s) = t$. For example, if χ_0 is the principal character modulo 3, then

$$L(s,\chi_0) = 1 + \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{8^s} + \cdots.$$

If χ_3 is the nonprincipal character modulo 3, then

$$L(s,\chi_3) = 1 - \frac{1}{2^s} + \frac{1}{4^s} - \frac{1}{5^s} + \frac{1}{7^s} - \frac{1}{8^s} + \cdots.$$

We shall prove that if χ_0 is the principal character modulo m, then $L(s,\chi_0)$ is analytic in the half-plane $\sigma > 1$, and if χ is a nonprincipal character modulo m, then $L(s,\chi)$ is analytic in the half-plane $\sigma > 0$ and, moreover, $L(1,\chi) \neq 0$. We shall see that this implies Dirichlet's theorem on primes in arithmetic progressions.

Theorem 10.3 *Let χ be a Dirichlet character modulo m, and let s be a complex number with $\Re(s) = \sigma > 1$. The function $L(s,\chi)$ is analytic and has the Euler product*

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

Moreover, $L(s,\chi) \neq 0$ and

$$\log L(s,\chi) = \sum_p \frac{\chi(p)}{p^s} + O(1). \tag{10.1}$$

Proof. Since

$$\left|\frac{\chi(n)}{n^s}\right| \leq \frac{1}{n^\sigma}$$

and

$$\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$$

converges for $\sigma > 1$, it follows that the series $L(s,\chi)$ converges uniformly and absolutely in the half-plane $\sigma \geq 1+\delta$ for every $\delta > 0$. Similarly, for every prime p, the series $\sum_{k=0}^{\infty} \chi(p^k)p^{-ks}$ converges uniformly and absolutely in the half-plane $\sigma > 1$, and

$$\sum_{k=0}^{\infty} \frac{\chi(p^k)}{p^{ks}} = \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$$

Since the character χ is completely multiplicative, the Fundamental Theorem of Arithmetic implies that

$$\prod_{p \leq x} \left(\sum_{k=0}^{\infty} \frac{\chi(p^k)}{p^{ks}} \right) = \sum_{n \in \mathcal{N}(x)} \frac{\chi(n)}{n^s},$$

where $\mathcal{N}(x)$ denotes the set of all positive integers n divisible only by primes $p \leq x$. In particular, if $n \leq x$ and p divides n, then $p \leq x$, and so $n \in \mathcal{N}(x)$.

For every $\varepsilon > 0$ there exists a number $x_0(\varepsilon)$ such that, if $x \geq x_0(\varepsilon)$, then

$$\sum_{n>x} \frac{1}{n^\sigma} < \varepsilon.$$

It follows that for $x \geq x_0(\varepsilon)$ we have

$$\begin{aligned}
\left| L(s,\chi) - \prod_{p \leq x} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \right| &= \left| \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} - \sum_{n \in \mathcal{N}(x)} \frac{\chi(n)}{n^s} \right| \\
&\leq \sum_{n>x} \left| \frac{\chi(n)}{n^s} \right| \\
&\leq \sum_{n>x} \frac{1}{n^\sigma} \\
&< \varepsilon,
\end{aligned}$$

and so the infinite product converges to the L-function, that is,

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

This product is called the Euler product of $L(s,\chi)$.

We shall prove directly that $L(s,\chi)$ is nonzero for $\sigma > 1$. Each factor of the Euler product is nonzero, since

$$\left|\frac{\chi(p)}{p^s}\right| \le \frac{1}{p^\sigma} < \frac{1}{2},$$

and so it suffices to prove that

$$\prod_{p>x_0}\left(1-\frac{\chi(p)}{p^s}\right)^{-1} \neq 0$$

for some number x_0. The inequality

$$\left|\sum_{k=1}^{\infty}\frac{\chi(p)}{p^{ks}}\right| \le \sum_{k=1}^{\infty}\frac{1}{p^{k\sigma}} = \frac{1}{p^\sigma - 1} < \frac{2}{p^\sigma}$$

implies that

$$\begin{aligned}\left|\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\right| &= \left|1+\sum_{k=1}^{\infty}\frac{\chi(p)}{p^{ks}}\right| \\ &\ge 1-\left|\sum_{k=1}^{\infty}\frac{\chi(p)}{p^{ks}}\right| \\ &> 1-\frac{2}{p^\sigma}.\end{aligned}$$

Choose x_0 such that

$$\sum_{p>x_0}\frac{2}{p^\sigma} < \frac{1}{2}.$$

It follows that for $x \ge x_0$ we have

$$\begin{aligned}\left|\prod_{x_0<p\le x}\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\right| &= \prod_{x_0<p\le x}\left|\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\right| \\ &\ge \prod_{x_0<p\le x}\left(1-\frac{2}{p^\sigma}\right) \\ &\ge 1-\sum_{x_0<p\le x}\frac{2}{p^\sigma} \\ &> \frac{1}{2},\end{aligned}$$

and so

$$\left|\prod_{p>x_0}\left(1-\frac{\chi(p)}{p^s}\right)^{-1}\right| \ge \frac{1}{2}.$$

Therefore,

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \neq 0.$$

For $|z| < 1$, the principal value of the logarithm has the power series

$$\log \frac{1}{1-z} = -\log(1-z) = \sum_{n=1}^{\infty} \frac{z^n}{n}.$$

Applying this to the Dirichlet L-function for $\sigma > 1$, we obtain

$$\begin{aligned}
\log L(s,\chi) &= \log \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \\
&= -\sum_p \log\left(1 - \frac{\chi(p)}{p^s}\right) \\
&= \sum_p \sum_{n=1}^{\infty} \frac{\chi(p^n)}{np^{ns}} \\
&= \sum_p \frac{\chi(p)}{p^s} + \sum_p \sum_{n=2}^{\infty} \frac{\chi(p^n)}{np^{ns}} \\
&= \sum_p \frac{\chi(p)}{p^s} + O(1),
\end{aligned}$$

since

$$\begin{aligned}
\left| \sum_p \sum_{n=2}^{\infty} \frac{\chi(p^n)}{np^{ns}} \right| &\leq \sum_p \sum_{n=2}^{\infty} \frac{1}{np^{n\sigma}} \\
&< \sum_p \sum_{n=2}^{\infty} \frac{1}{p^n} \\
&= \sum_p \frac{1}{p(p-1)} \\
&\ll 1.
\end{aligned}$$

This completes the proof. □

For example, if χ_0 and χ_3 are the principal and nonprincipal characters modulo 3, respectively, then

$$L(s,\chi_0) = \prod_{p \geq 3} \left(1 - p^{-s}\right)^{-1}$$

and

$$L(s,\chi_3) = \prod_{p\equiv 1 \pmod 3} \left(1-p^{-s}\right)^{-1} \prod_{p\equiv 2 \pmod 3} \left(1+p^{-s}\right)^{-1}.$$

Theorem 10.4 *Let χ be a nonprincipal character modulo m. The Dirichlet L-function $L(s,\chi)$ is analytic in the half-plane $\sigma > 0$. Let K be a compact set in the half-plane $\sigma > 0$. For $s \in K$ and $x \geq 1$,*

$$L(s,\chi) = \sum_{n\leq x} \frac{\chi(n)}{n^s} + O\left(x^{-\sigma}\right), \tag{10.2}$$

where the implied constant depends on m and K.

Proof. To prove that $L(s,\chi)$ is analytic in $\sigma > 0$, it suffices to prove that the series $\sum_{n=1}^{\infty} \chi(n)n^{-s}$ converges uniformly on every compact subset of the right half-plane $\sigma > 0$.

Let K be a compact subset of the right half-plane. There exist positive constants δ and Δ such that $\sigma \geq \delta$ and $|s| \leq \Delta$ for every $s = \sigma + it \in K$. We use partial summation (Theorem 6.8) with

$$f(n) = \chi(n)$$

and

$$g(t) = \frac{1}{t^s}.$$

By Exercise 6 in Section 10.1, $F(t) = \sum_{n\leq t} \chi(n) \ll 1$ and

$$\begin{aligned}
\sum_{N<n\leq M} \frac{\chi(n)}{n^s} &= \sum_{N<n\leq M} f(n)g(n) \\
&= F(M)g(M) - F(N)g(N) - \int_N^M F(t)g'(t)dt \\
&= \frac{F(M)}{M^s} - \frac{F(N)}{N^s} + s\int_N^M \frac{F(t)}{t^{s+1}}dt \\
&\ll \frac{1}{M^\sigma} + \frac{1}{N^\sigma} + |s| \int_N^M \frac{1}{t^{\sigma+1}}dt \\
&\ll \frac{1}{N^\sigma} + \frac{|s|}{\sigma N^\sigma} \\
&\ll \left(1+\frac{\Delta}{\delta}\right)\frac{1}{N^\sigma} \\
&\ll \frac{1}{N^\delta},
\end{aligned}$$

where the implied constants depend on the modulus m and the compact set K. It follows that the partial sums of the series $L(s,\chi)$ are uniformly

Cauchy on K, and so $L(s,\chi)$ converges uniformly on K and is analytic in the right half-plane.

Since

$$\sum_{N<n\leq M} \frac{\chi(n)}{n^s} \ll \frac{1}{N^\sigma}$$

for all $M > N$, it follows that

$$L(s,\chi) - \sum_{n=1}^{N} \frac{\chi(n)}{n^s} = \sum_{n=N}^{\infty} \frac{\chi(n)}{n^s} \ll \frac{1}{N^\sigma}.$$

This completes the proof. □

The analytic nature of Dirichlet L-functions is different for principal and nonprincipal characters. In the special case where χ_0 is the principal character modulo 1, we have $\chi_0(n) = 1$ for all integers n, and the Dirichlet L-function $L(s,\chi_0)$ for $\sigma > 1$ is the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Theorem 10.5 *Let χ_0 be the principal character modulo m. For $\sigma > 1$,*

$$L(s,\chi_0) = \zeta(s) \prod_{p|m} \left(1 - \frac{1}{p^s}\right)$$

and

$$\lim_{\sigma \to 1^+} (\sigma - 1) L(\sigma,\chi_0) = \prod_{p|m} \left(1 - \frac{1}{p}\right).$$

For $1 < \sigma < 2$,

$$\log L(\sigma,\chi_0) = \log\left(\frac{1}{\sigma - 1}\right) + O(1).$$

Proof. The Riemann zeta function is not analytic at $s = 1$, since for $\sigma > 1$ and $n \geq 1$ we have

$$\int_n^{n+1} \frac{dx}{x^\sigma} < \frac{1}{n^\sigma} < \int_{n-1}^{n} \frac{dx}{x^\sigma},$$

and so

$$0 < \frac{1}{\sigma - 1} = \int_1^\infty \frac{dx}{x^\sigma} < \zeta(\sigma) < 1 + \int_1^\infty \frac{dx}{x^\sigma} = \frac{\sigma}{\sigma - 1}.$$

Therefore,

$$1 < (\sigma - 1)\zeta(\sigma) < \sigma$$

and

$$\lim_{\sigma\to 1^+}(\sigma-1)\zeta(\sigma)=1. \tag{10.3}$$

If $1<\sigma<2$, then

$$\log\left(\frac{1}{\sigma-1}\right)<\log\zeta(\sigma)<\log\left(\frac{1}{\sigma-1}\right)+\log\sigma<\log\left(\frac{1}{\sigma-1}\right)+\log 2,$$

and so

$$\log\zeta(\sigma)=\log\left(\frac{1}{\sigma-1}\right)+O(1). \tag{10.4}$$

If χ_0 is the principal character modulo m, then

$$\begin{aligned}
L(s,\chi_0) &= \prod_p\left(1-\frac{\chi_0(p)}{p^s}\right)^{-1}\\
&= \prod_{(p,m)=1}\left(1-\frac{1}{p^s}\right)^{-1}\\
&= \prod_p\left(1-\frac{1}{p^s}\right)^{-1}\prod_{p|m}\left(1-\frac{1}{p^s}\right)\\
&= \zeta(s)\prod_{p|m}\left(1-\frac{1}{p^s}\right).
\end{aligned}$$

Let $1<\sigma<2$. Then

$$(\sigma-1)L(\sigma,\chi_0)=(\sigma-1)\zeta(\sigma)\prod_{p|m}\left(1-\frac{1}{p^\sigma}\right),$$

and (10.3) implies that

$$\lim_{\sigma\to 1^+}(\sigma-1)L(\sigma,\chi_0)=\prod_{p|m}\left(1-\frac{1}{p}\right).$$

Moreover,

$$\begin{aligned}
\log L(\sigma,\chi_0) &= \log\zeta(\sigma)+\log\prod_{p|m}\left(1-\frac{1}{p^\sigma}\right)\\
&= \log\left(\frac{1}{\sigma-1}\right)+O(1),
\end{aligned}$$

by (10.4). □

Exercises

1. Compute the four Dirichlet L-functions modulo 8.

2. A *Dirichlet series* is a function of the form

$$F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

where $\{a_n\}_{n=1}^{\infty}$ is a sequence of complex numbers. Prove that if $a_n = O(n^{\alpha})$, then the series $F(s)$ converges in the half-plane $\sigma > 1+\alpha$ and uniformly in the half-plane $\sigma \geq 1+\alpha+\delta$ for every $\delta > 0$.

3. A *Dirichlet polynomial* is a function of the form

$$F(s) = \sum_{n=1}^{N} \frac{a_n}{n^s},$$

where $\{a_n\}_{n=1}^{N}$ is a finite sequence of complex numbers. Find the zeros of the Dirichlet polynomial

$$\sum_{d|m} \frac{\mu(d)}{d^s} = \prod_{p|m} \left(1 - p^{-s}\right).$$

4. Let χ_0 be the principal character modulo 3, and let χ_3 be the non-principal character modulo 3. Prove that

$$L(s,\chi_0) + L(s,\chi_3) = 2 \sum_{\substack{n=1 \\ n\equiv 1 \pmod 3}}^{\infty} \frac{1}{n^3}.$$

5. Let $m \geq 4$, and let $\widehat{G}$ be the group of Dirichlet characters modulo m. Prove that

$$\sum_{\chi\in\widehat{G}} L(s,\chi) = \varphi(m) \sum_{\substack{n=1 \\ n\equiv 1 \pmod m}} \frac{1}{n^s}.$$

6. Let k and n be positive integers such that k divides n, let χ^* be a Dirichlet character modulo k, and let χ be the Dirichlet character modulo m induced by χ^*. Prove that

$$L(s,\chi) = L(s,\chi^*) \prod_{p|m} \left(1 - \frac{\chi^*(p)}{p^s}\right).$$

7. Let

$$f(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}. \tag{10.5}$$

Prove that

(a) $f(s)$ is analytic in the half-plane $\sigma > 0$,

(b) $0 < f(\sigma) < 1$ for $\sigma > 0$.

8. Let
$$g(s) = 1 - 2^{1-s}. \tag{10.6}$$
Prove that

(a) $g(s)$ is analytic in the entire complex plane.

(b) $g(s) = 0$ if and only if $s = 1 - 2\pi ik/\log 2$ for $k \in \mathbf{Z}$.

(c) $g'(1 - 2\pi ik/\log 2) = \log 2$.

(d) $g(\sigma) < 0$ for $0 < \sigma < 1$.

(e) $\left(1 - 2^{1-s}\right)^{-1}$ is meromorphic in the complex plane except for simple poles at $s = 1 - 2\pi ik/\log 2$ with residues $1/\log 2$.

9. Define the functions $f(s)$ and $g(s)$ by (10.5) and (10.6), respectively. Prove that for $\sigma > 1$,
$$f(s) = g(s)\zeta(s),$$
or, equivalently,
$$\zeta(s) = \left(1 - 2^{1-s}\right)^{-1} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}.$$
Show that the right side of this equation is meromorphic in the half-plane $\sigma > 0$. This determines the meromorphic continuation of the Riemann zeta function to the half-plane $\sigma > 0$. Prove that
$$\zeta(\sigma) < 0 \qquad \text{for } 0 < \sigma < 1.$$
Use (10.3) to prove that
$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2.$$

10.3 Primes Modulo 4

In this section we show that there are infinitely many primes p such that $p \equiv 1 \pmod 4$, and also infinitely many primes p such that $p \equiv 3 \pmod 4$. This is Dirichlet's theorem for modulus 4. The proof is easier than the general case, and shows clearly the use of Dirichlet characters and Dirichlet L-functions.

There are two Dirichlet characters modulo 4. Let χ_0 be the principal Dirichlet character. Then

$$\chi_0(n) = \begin{cases} 1 & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

The L-function $L(s, \chi_0)$ converges in the half-plane $\sigma > 1$, where

$$\begin{aligned} L(s, \chi_0) &= \sum_{n=1}^{\infty} \frac{1}{(2n-1)^s} = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \cdots \\ &= \prod_{p \neq 2} \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \left(1 - \frac{1}{2^s}\right) \zeta(s), \end{aligned}$$

but the infinite series

$$L(1, \chi_0) = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots$$

diverges.

Let χ_4 be the nonprincipal character modulo 4. Then

$$\chi_4(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod 4, \\ -1 & \text{if } n \equiv 3 \pmod 4, \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

The L-function $L(s, \chi_0)$ converges in the half-plane $\sigma > 0$, where

$$\begin{aligned} L(s, \chi_4) &= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^s} = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots \\ &= \prod_{p \equiv 1 \pmod 4} \left(1 - \frac{1}{p^s}\right)^{-1} \prod_{p \equiv 3 \pmod 4} \left(1 + \frac{1}{p^s}\right)^{-1}. \end{aligned}$$

The infinite series

$$L(1, \chi_4) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

converges, and $L(1, \chi_4) > 0$. Indeed,

$$\begin{aligned} L(1, \chi_4) &= \left(1 - \frac{1}{3}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) + \left(\frac{1}{9} - \frac{1}{11}\right) + \cdots \\ &> 0.744 \end{aligned}$$

and

$$\begin{aligned} L(1,\chi_4) &= 1-\left(\frac{1}{3}-\frac{1}{5}\right)-\left(\frac{1}{7}-\frac{1}{9}\right)-\cdots \\ &< 0.835. \end{aligned}$$

(Using the power series expansion of the inverse tangent, one can prove Leibniz's formula $L(1,\chi_4) = \pi/4 = 0.785\ldots$.)

Theorem 10.6 *For* $1 < \sigma < 2$,

$$\sum_{p \equiv 1 \pmod 4} \frac{1}{p^\sigma} = \frac{1}{2}\log\frac{1}{\sigma-1} + O(1)$$

and

$$\sum_{p \equiv 3 \pmod 4} \frac{1}{p^\sigma} = \frac{1}{2}\log\frac{1}{\sigma-1} + O(1).$$

In particular, there exist infinitely many primes $p \equiv 1 \pmod 4$ *and infinitely many primes* $p \equiv 3 \pmod 4$.

Proof. Since $L(s,\chi_4)$ is continuous for $\sigma > 0$, it follows that

$$\log L(\sigma,\chi_4) = O(1) \qquad \text{for } 1 \leq \sigma \leq 2.$$

Let $1 < \sigma < 2$. By (10.1) of Theorem 10.3, we have

$$\log L(\sigma,\chi_0) = \sum_{p\geq 3}\frac{1}{p^\sigma} + O(1)$$

and

$$\log L(\sigma,\chi_4) = \sum_{p\geq 3}\frac{(-1)^{(p-1)/2}}{p^\sigma} + O(1).$$

Therefore,

$$\begin{aligned} \sum_{p \equiv 1 \pmod 4} \frac{1}{p^\sigma} &= \frac{1}{2}\left(\log L(\sigma,\chi_0) + \log L(\sigma,\chi_4)\right) + O(1) \\ &= \frac{1}{2}\log L(\sigma,\chi_0) + O(1) \\ &= \frac{1}{2}\log\frac{1}{\sigma-1} + O(1), \end{aligned}$$

by Theorem 10.5. Since

$$\lim_{\sigma\to 1^+}\log\frac{1}{\sigma-1} = \infty,$$

it follows that there exist infinitely many primes congruent to 1 modulo 4. Similarly,

$$\sum_{p\equiv 3 \pmod 4} \frac{1}{p^\sigma} = \frac{1}{2}\log\frac{1}{\sigma-1} + O(1),$$

and there exist infinitely many primes congruent to 3 modulo 4. □

Exercises

1. Let χ_0 be the principal Dirichlet character modulo 6, and let χ_6 be the nonprincipal Dirichlet character modulo 6. Prove that

$$\sum_{p\equiv 1 \pmod 6} \frac{1}{p^\sigma} = \frac{1}{2}\left(\log L(\sigma,\chi_0) + \log L(\sigma,\chi_6)\right) + O(1)$$

   and

$$\sum_{p\equiv 5 \pmod 6} \frac{1}{p^\sigma} = \frac{1}{2}\left(\log L(\sigma,\chi_0) - \log L(\sigma,\chi_6)\right) + O(1).$$

2. Prove that there exist infinitely many primes $p \equiv 1 \pmod 6$ and infinitely many primes $p \equiv 5 \pmod 6$.

3. Compute $L(1,\chi_6)$ with an error of at most 0.01.

10.4 The Nonvanishing of $L(1,\chi)$

In this section we prove that $L(1,\chi) \neq 0$ for every nonprincipal character χ.

Lemma 10.1 *Let χ_0 be the principal character modulo m. Then*

$$\sum_{n\leq x} \frac{\chi_0(n)\Lambda(n)}{n} = \log x + O(1).$$

Proof. Observe that

$$\begin{aligned}
\sum_{\substack{n\leq x\\ (n,m)>1}} \frac{\Lambda(n)}{n} &= \sum_{p\mid m}\sum_{\substack{p^k\leq x\\ k\geq 1}} \frac{\Lambda(p^k)}{p^k} \\
&< \sum_{p\mid m}\sum_{k=1}^{\infty} \frac{\log p}{p^k}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{p|m} \frac{\log p}{p-1} \\
&= O(1).
\end{aligned}
$$

By Mertens's theorem (Theorem 8.5), we have

$$
\begin{aligned}
\sum_{n\leq x} \frac{\chi_0(n)\Lambda(n)}{n} &= \sum_{\substack{n\leq x\\(n,m)=1}} \frac{\Lambda(n)}{n} \\
&= \sum_{n\leq x} \frac{\Lambda(n)}{n} - \sum_{\substack{n\leq x\\(n,m)>1}} \frac{\Lambda(n)}{n} \\
&= \log x + O(1).
\end{aligned}
$$

This completes the proof. □

Lemma 10.2 *Let χ be a nonprincipal character modulo m. If $L(1,\chi) \neq 0$, then*

$$
\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} = O(1).
$$

Proof. Recall that $F(t) = \sum_{k\leq t} \chi(k) \ll 1$ (Exercise 6 in Section 10.1). By partial summation, we have

$$
\begin{aligned}
\sum_{k\leq x} \frac{\chi(k)\log k}{k} &= \frac{F(x)\log x}{x} - \int_1^x \frac{F(t)(1-\log t)}{t^2} dt \\
&\ll \frac{\log x}{x} + \int_1^\infty \frac{1+\log t}{t^2} dt \\
&\ll 1.
\end{aligned}
$$

By Theorem 10.4, we have

$$
L(1,\chi) = \sum_{d\leq x/n} \frac{\chi(d)}{d} + O\left(\frac{n}{x}\right).
$$

Using the identity $\log k = \sum_{n|k} \Lambda(n)$, we obtain

$$
\begin{aligned}
\sum_{k\leq x} \frac{\chi(k)\log k}{k} &= \sum_{k\leq x} \frac{\chi(k)}{k} \sum_{n|k} \Lambda(n) \\
&= \sum_{nd\leq x} \frac{\chi(nd)\Lambda(n)}{nd}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} \sum_{d\leq x/n} \frac{\chi(d)}{d} \\
&= \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} \left(L(1,\chi) + O\left(\frac{n}{x}\right)\right) \\
&= L(1,\chi) \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} + \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} O\left(\frac{n}{x}\right) \\
&= L(1,\chi) \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} + O(1),
\end{aligned}
$$

since

$$
\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} O\left(\frac{n}{x}\right) \ll \frac{1}{x} \sum_{n\leq x} \Lambda(n) = \frac{\psi(x)}{x} \ll 1
$$

by Chebyshev's theorem (Theorem 8.2). Therefore,

$$
L(1,\chi) \sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} = O(1).
$$

If $L(1,\chi) \neq 0$, then

$$
\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} = O(1).
$$

This completes the proof. □

Lemma 10.3 *Let χ be a nonprincipal character modulo m. If $L(1,\chi) = 0$, then*

$$
\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} = -\log x + O(1).
$$

Proof. Since

$$
\Lambda(n) = -\sum_{d|n} \mu(d) \log d,
$$

we have

$$
\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} = -\sum_{n\leq x} \frac{\chi(n)}{n} \sum_{d|n} \mu(d) \log d.
$$

From the identity

$$
\log x = \sum_{n\leq x} \frac{\chi(n)}{n} \sum_{d|n} \mu(d) \log x,
$$

we have

$$\begin{aligned}
\log x + \sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} \\
&= \sum_{n \le x} \frac{\chi(n)}{n} \sum_{d|n} \mu(d) \log \frac{x}{d} \\
&= \sum_{dk \le x} \frac{\chi(dk)\mu(d)}{dk} \log \frac{x}{d} \\
&= \sum_{d \le x} \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} \sum_{k \le x/d} \frac{\chi(k)}{k} \\
&= \sum_{d \le x} \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} \left(L(1,\chi) + O\left(\frac{d}{x}\right)\right) \\
&= L(1,\chi) \sum_{d \le x} \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} + \sum_{d \le x} O\left(\frac{d}{x}\right) \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} \\
&= L(1,\chi) \sum_{d \le x} \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} + O(1),
\end{aligned}$$

since

$$\sum_{d \le x} O\left(\frac{d}{x}\right) \frac{\chi(d)\mu(d)}{d} \log \frac{x}{d} \ll \frac{1}{x} \sum_{d \le x} \log \frac{x}{d} \ll 1$$

by Theorem 6.4. If $L(1,\chi) = 0$, then

$$\sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n} = -\log x + O(1).$$

This completes the proof. □

Theorem 10.7 *Let χ be a complex character modulo m. Then $L(1,\chi) \neq 0$.*

Proof. Let N denote the number of nonprincipal characters modulo m such that $L(1,\chi) = 0$. We shall prove that $N = 0$ or 1. By Lemmas 10.1, 10.2, and 10.3, and the orthogonality relations for Dirichlet characters (Theorem 10.1), we have

$$\begin{aligned}
\varphi(m) \sum_{\substack{n \le x \\ n \equiv 1 \pmod m}} \frac{\Lambda(n)}{n} &= \sum_{n \le x} \frac{\Lambda(n)}{n} \sum_{\chi \pmod m} \chi(n) \\
&= \sum_{\chi \pmod m} \sum_{n \le x} \frac{\chi(n)\Lambda(n)}{n}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{n\leq x} \frac{\chi_0(n)\Lambda(n)}{n} + \sum_{\chi\neq\chi_0}\sum_{n\leq x} \frac{\chi(n)\Lambda(n)}{n} \\
&= \log x - N\log x + O(1) \\
&= (1-N)\log x + O(1).
\end{aligned}
$$

Since $\Lambda(n)/n \geq 0$ for all $n \geq 1$, it follows that both sides of this equation are nonnegative for large x, and so $N \leq 1$. Therefore, $L(1,\chi) = 0$ for at most one nonprincipal character χ.

If χ is a complex character modulo m, then $\overline{\chi}$ is also a complex character and $\chi \neq \overline{\chi}$. We have

$$\overline{L(1,\chi)} = \overline{\sum_{n=1}^{\infty} \frac{\chi(n)}{n}} = \sum_{n=1}^{\infty} \frac{\overline{\chi}(n)}{n} = L(1,\overline{\chi}),$$

and so $L(1,\chi) = 0$ if and only if $L(1,\overline{\chi}) = 0$. Since $N \leq 1$, we must have $L(1,\chi) \neq 0$ for every complex character χ. This completes the proof. □

Theorem 10.8 *Let χ be a real nonprincipal character modulo m. Then $L(1,\chi) \neq 0$.*

Proof. Since the character χ is real, we have $\chi(n) = \pm 1$ for every integer n. Consequently, for every prime power p^r,

$$\sum_{j=0}^{r} \chi(p^j) = 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p)^r \geq 0$$

and

$$\sum_{j=0}^{r} \chi(p^j) \geq 1 \quad \text{if } r \text{ is even.}$$

The character χ is multiplicative, and so the convolution

$$t(k) = 1 * \chi(k) = \sum_{d\mid k} \chi(d)$$

is also a multiplicative function. It follows that

$$t(k) = \prod_{p^r\| k} t(p^r) = \prod_{p^r\| k} \sum_{j=0}^{r} \chi(p^j) \geq 0$$

and

$$t(k) \geq 1 \qquad \text{if } k = m^2 \text{ is a square.}$$

Using the asymptotic formula in Theorem 6.9 for the partial sums of the harmonic series, we obtain for large x the lower bound

$$\begin{aligned} T(x) &= \sum_{k \leq x} \frac{t(k)}{k^{1/2}} \geq \sum_{m^2 \leq x} \frac{t(m^2)}{m} \\ &\geq \sum_{m \leq x^{1/2}} \frac{1}{m} > \frac{\log x}{2}. \end{aligned}$$

Applying the L-function estimate (10.2) in Theorem 10.4 with $s = 1$ and $s = 1/2$, we have

$$\sum_{n \leq x} \frac{\chi(n)}{n} = L(1, \chi) + O\left(x^{-1}\right)$$

and

$$\sum_{n \leq x} \frac{\chi(n)}{n^{1/2}} = L(1/2, \chi) + O\left(x^{-1/2}\right).$$

Let $x \geq 1$ and $y = x^{1/2}$. By Exercise 6, the set of all lattice points (n, d) such that n and d are positive and $nd \leq x$ can be partitioned into two disjoint sets as follows: The first set consists of all lattice points (n, d) such that $1 \leq n \leq x^{1/2}$ and $1 \leq d \leq x/n$, and the second set consists of all lattice points (n, d) such that $1 \leq d < x^{1/2}$ and $x^{1/2} < n \leq x/d$. If $d = x^{1/2}$, then $x/d = x^{1/2}$ and there is no integer n such that $x^{1/2} < n \leq x/d$. Therefore, the second set can also be described as the set of all lattice points (n, d) such that $1 \leq d \leq x^{1/2}$ and $x^{1/2} < n \leq x/d$. We have

$$\begin{aligned} T(x) &= \sum_{k \leq x} \frac{t(k)}{k^{1/2}} \\ &= \sum_{k \leq x} \frac{1}{k^{1/2}} \sum_{n | k} \chi(n) \\ &= \sum_{nd \leq x} \frac{\chi(n)}{(nd)^{1/2}} \\ &= \sum_{n \leq x^{1/2}} \sum_{d \leq x/n} \frac{\chi(n)}{(nd)^{1/2}} + \sum_{d \leq x^{1/2}} \sum_{x^{1/2} < n \leq x/d} \frac{\chi(n)}{(nd)^{1/2}} \\ &= \sum_{n \leq x^{1/2}} \frac{\chi(n)}{n^{1/2}} \sum_{d \leq x/n} \frac{1}{d^{1/2}} + \sum_{d \leq x^{1/2}} \frac{1}{d^{1/2}} \sum_{x^{1/2} < n \leq x/d} \frac{\chi(n)}{n^{1/2}}. \end{aligned}$$

We shall estimate these sums separately. By Exercise 7,

$$\sum_{d \leq x} \frac{1}{d^{1/2}} = 2x^{1/2} - c + O\left(x^{-1/2}\right).$$

The first sum is

$$\begin{aligned}
&\sum_{n\le x^{1/2}} \frac{\chi(n)}{n^{1/2}} \sum_{d\le x/n} \frac{1}{d^{1/2}} \\
&= \sum_{n\le x^{1/2}} \frac{\chi(n)}{n^{1/2}} \left(\frac{2x^{1/2}}{n^{1/2}} - c + O\left(\frac{n^{1/2}}{x^{1/2}}\right)\right) \\
&= 2x^{1/2} \sum_{n\le x^{1/2}} \frac{\chi(n)}{n} - c \sum_{n\le x^{1/2}} \frac{\chi(n)}{n^{1/2}} + O\left(\sum_{n\le x^{1/2}} \frac{1}{x^{1/2}} \right) \\
&= 2x^{1/2}\left(L(1,\chi) + O\left(x^{-1/2}\right)\right) - cL(1/2,\chi) + O\left(x^{-1/4}\right) + O(1) \\
&= 2L(1,\chi)x^{1/2} + O(1).
\end{aligned}$$

The second sum is

$$\begin{aligned}
&\sum_{d\le x^{1/2}} \frac{1}{d^{1/2}} \sum_{x^{1/2}<n\le x/d} \frac{\chi(n)}{n^{1/2}} \\
&= \sum_{d\le x^{1/2}} \frac{1}{d^{1/2}} \left(\left(L(1/2,\chi) + O\left(\frac{d^{1/2}}{x^{1/2}}\right)\right) - \left(L(1/2,\chi) + O\left(x^{-1/4}\right)\right)\right) \\
&= \sum_{d\le x^{1/2}} \frac{1}{d^{1/2}} \left(O\left(\frac{d^{1/2}}{x^{1/2}}\right) + O\left(x^{-1/4}\right)\right) \\
&\ll 1 + \frac{1}{x^{1/4}} \sum_{d\le x^{1/2}} \frac{1}{d^{1/2}} \\
&\ll 1 + \frac{1}{x^{1/4}}\left(x^{1/4}+1\right) \\
&\ll 1.
\end{aligned}$$

Therefore,

$$T(x) = 2L(1,\chi)x^{1/2} + O(1).$$

However, we also have

$$T(x) > \frac{\log x}{2}$$

for sufficiently large x, which is impossible if $L(1,\chi) = 0$. Therefore, $L(1,\chi) \neq 0$ for all nonprincipal real characters χ. □

We can now prove Dirichlet's theorem.

Theorem 10.9 (Dirichlet) *Let m and a be relatively prime positive integers. For $1 < \sigma < 2$,*

$$\sum_{p\equiv a \pmod m} \frac{1}{p^\sigma} = \frac{1}{\varphi(m)} \log\left(\frac{1}{\sigma-1}\right) + O(1)$$

In particular, there exist infinitely primes p such that $p \equiv a \pmod{m}$.

Proof. Let $1 < \sigma < 2$. Using the orthogonality relations for Dirichlet characters (Theorem 10.2) and the estimate (10.1) for $\log L(s, \chi)$ from Theorem 10.3, we obtain

$$\begin{aligned} \sum_{\chi \pmod{m}} \overline{\chi}(a) \log L(\sigma, \chi) &= \sum_{\chi \pmod{m}} \sum_{p} \frac{\overline{\chi}(a)\chi(p)}{p^{\sigma}} + O(1) \\ &= \sum_{p} \frac{1}{p^{\sigma}} \sum_{\chi \pmod{m}} \overline{\chi}(a)\chi(p) + O(1) \\ &= \varphi(m) \sum_{p \equiv a \pmod{m}} \frac{1}{p^{\sigma}} + O(1). \end{aligned}$$

By Theorem 10.5, the term on the left corresponding to the principal character χ_0 is

$$\overline{\chi_0}(a) \log L(\sigma, \chi_0) = \log\left(\frac{1}{\sigma - 1}\right) + O(1),$$

and so

$$\varphi(m) \sum_{p \equiv a \pmod{m}} \frac{1}{p^{\sigma}} = \log\left(\frac{1}{\sigma - 1}\right) + \sum_{\chi \neq \chi_0} \overline{\chi}(a) \log L(\sigma, \chi) + O(1).$$

If χ is a nonprincipal character modulo m, then $L(1, \chi) \neq 0$ by Theorem 10.7 and Theorem 10.8, and so $\log L(\sigma, \chi) = O(1)$ for $1 \leq \sigma \leq 2$. This proves that

$$\sum_{p \equiv a \pmod{m}} \frac{1}{p^{\sigma}} = \frac{1}{\varphi(m)} \log\left(\frac{1}{\sigma - 1}\right) + O(1).$$

Therefore, the series $\sum_{p \equiv a \pmod{m}} p^{-\sigma}$ diverges as $\sigma \to 1^{+}$, and so it must have infinitely many terms, that is, there must exist infinitely primes p such that $p \equiv a \pmod{m}$. This completes the proof of Dirichlet's theorem. □

Finally, we obtain a generalization of Mertens's theorem (Theorem 8.5) to sums of $\Lambda(n)/n$ over an arithmetic progression.

Theorem 10.10 *Let $m \geq 1$ and a be relatively prime integers. Then*

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod{m}}} \frac{\Lambda(n)}{n} = \frac{\log x}{\varphi(m)} + O(1).$$

Proof. For the principal character χ_0 we have

$$\sum_{n \leq x} \frac{\chi_0(n)\Lambda(n)}{n} = \log x + O(1)$$

by Lemma 10.1. For every nonprincipal character χ modulo m, we have $L(1,\chi) \neq 0$ by Theorems 10.7 and 10.8, and so

$$\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = O(1)$$

by Lemma 10.2. Since $\chi_0(a) = 1$, it follows that

$$\sum_{\chi \pmod{m}} \overline{\chi}(a) \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \overline{\chi}(a) \log x + O(1) = \log x + O(1).$$

On the other hand, by Theorem 10.2,

$$\begin{aligned} \sum_{\chi \pmod{m}} \overline{\chi(a)} \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} &= \sum_{n \leq x} \frac{\Lambda(n)}{n} \sum_{\chi \pmod{m}} \overline{\chi(a)}\chi(n) \\ &= \varphi(m) \sum_{\substack{n \leq x \\ n \equiv a \pmod{m}}} \frac{\Lambda(n)}{n}. \end{aligned}$$

This completes the proof. □

Exercises

1. Let χ_4 be the nonprincipal character modulo 4. Prove that

$$L(1,\chi_4) = 2\sum_{n=1}^{\infty} \frac{1}{(4n-2)^2 - 1} = 1 - 2\sum_{n=2}^{\infty} \frac{1}{16n^2 - 1}.$$

2. Let χ_3 be the nonprincipal character modulo 3. Prove that

$$L(1,\chi_3) = 2\sum_{n=0}^{\infty} \frac{1}{(3n+1)(3n+2)}.$$

3. Let χ be the Dirichlet character modulo 8 defined by $\chi(3) = \chi(5) = -1$. Show that

$$L(1,\chi) = 2\sum_{k=0}^{\infty} \frac{85k+32}{(8k+1)(8k+3)(8k+5)(8k+7)}.$$

4. Let χ_1 be the real primitive character modulo 5. Prove that $L(1,\chi) > 0$. Let χ_2 be the complex character modulo 5 defined by $\chi_2(2) = i$. Prove that the real and imaginary parts of $L(1,\chi_2)$ are positive.

5. Let m and a be relatively prime positive integers. Prove that

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod{m}}} \frac{\log p}{p} = \frac{\log x}{\varphi(m)} + O(1).$$

6. Prove that the set of all lattice points (n, d) such n and d are positive and $nd \leq x$ can be partitioned into two disjoint sets as follows: The first set consists of all lattice points (n, d) such that $1 \leq n \leq y$ and $1 \leq d \leq x/n$, and the second set consists of all lattice points (n, d) such that $1 \leq d < x/y$ and $y < n \leq x/d$.

7. Compute the constant c such that

$$\sum_{d \leq x} \frac{1}{d^{1/2}} = 2x^{1/2} - c + O(x^{-1/2}).$$

Hint. Partial summation.

10.5 Notes

Our proof of Dirichlet's theorem is "elementary" in the sense that it does not use complex analysis. Selberg [127] gave a different proof that is, he wrote, "more elementary in the respect that we do not use the complex characters mod k, and also in that we consider only finite sums."

Let m and a be relatively prime positive integers. We denote by $\pi(x; m, a)$ the number of prime numbers $p \leq x$ such that $p \equiv a \pmod{m}$. By the prime number theorem,

$$\pi(x) = \sum_{\substack{a=1 \\ (a,m)=1}}^{m} \pi(x; m, a) + \sum_{p \mid m} 1 \sim x.$$

The *prime number theorem for arithmetic progressions* states that for every integer $m \geq 3$ the prime numbers are uniformly distributed in the $\varphi(m)$ congruence classes relatively prime to m, that is, if $(a, m) = (b, m) = 1$, then

$$\pi(x; m, a) \sim \pi(x; m, b).$$

Equivalently, if $(a, m) = 1$, then

$$\pi(x; m, a) \sim \frac{x}{\varphi(m) \log x}.$$

Selberg [129] also gave an elementary proof of this result. Granville [39] reviews elementary proofs of the prime number theorem for arithmetic progressions. For an analytic proof, see Davenport [21].

For moduli $m \geq 3$ we can describe the *comparative prime number race* as follows. There are $\varphi(m)$ runners, one for each congruence class a relatively prime to m. For every positive integer x, the position of runner $a \pmod{m}$ at time x is $\pi(x; m, a)$. A runner *wins the mod m race* if it is eventually ahead of all the others. Does some congruence class win, or does the lead oscillate infinitely often between some or all of the competitors? In the case $m = 4$, Littlewood [94, 54] proved that $\pi(x; 4, 1) - \pi(x; 4, 3)$ changes sign infinitely often, so no class wins the "mod 4" race. More generally, we can ask the following question: Is it true that for every permutation $a_1, \ldots, a_{\varphi(m)}$ of the $\varphi(m)$ congruence classes relatively prime to m, we have

$$\pi(x; m, a_1) < \pi(x; m, a_2) < \cdots < \pi(x; m, a_{\varphi(m)})$$

for infinitely many integers x? This is an open problem in comparative prime number theory. For some results on this topic, see Turán [144].

In the Notes at the end of Chapter 9, we stated the Riemann hypothesis in the form

$$\pi(x) = \mathrm{li}(x) + O\left(x^{1/2+\varepsilon}\right)$$

for every $\varepsilon > 0$. In Exercise 9 of Section 10.2 we constructed the meromorphic continuation of the Riemann zeta function to the half-plane $\sigma > 0$. We can now state the Riemann hypothesis in its usual form: If $\zeta(s) = 0$ with $s = \sigma + it$ and $\sigma > 0$, then $\sigma = 1/2$.

Part III

Three Problems in Additive Number Theory

11
Waring's Problem

11.1 Sums of Powers

Lagrange proved that every number is the sum of four squares. This means that for every nonnegative integer n there exist nonnegative integers x_1, x_2, x_3, x_4 such that

$$n = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

Similarly, Wieferich proved that every number is the sum of nine cubes, that is, for every nonnegative integer n there exist nonnegative integers $x_1, \ldots, x_9$ such that

$$n = x_1^3 + x_2^3 + \cdots + x_9^3.$$

These are special cases of *Waring's problem*, one of the most famous problems in number theory. Waring's problem states that for every integer $k \geq 2$ there exists a number h such that every nonnegative integer can be written as the sum of exactly h kth powers. The smallest such integer h is usually denoted by $g(k)$. Since 7 cannot be written as the sum of three squares, and 23 cannot be written as the sum of 8 cubes, we can restate Lagrange's theorem as $g(2) = 4$, and Wieferich's theorem as $g(3) = 9$.

In 1909, the German mathematician David Hilbert proved Waring's problem for all exponents k. The British mathematicians G. H. Hardy and J. E. Littlewood subsequently devised a different proof, and their method was simplified and improved by the Soviet mathematician I. M. Vinogradov. These proofs involve sophisticated techniques of real and complex analysis, even though the statement of Waring's problem is purely arithmetic. In 1943, another Soviet mathematician, Yu. V. Linnik, devised a proof of

Waring's problem that uses only elementary number theory. In this and the following chapter we give Linnik's proof of Waring's problem.

There is a natural generalization of Waring's problem to polynomials. Let $f(x)$ be a polynomial of degree k that is *integer-valued*, that is, $f(x)$ is an integer for every integer x. Every polynomial with integer coefficients is integer-valued. There are also polynomials with rational coefficients that are integer-valued. For example, the binomial polynomial

$$b_k(x) = \binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}$$

is integer-valued, and every integral linear combination of binomial polynomials is integer-valued. Moreover, every integer-valued polynomial $f(x)$ of degree k can be expressed uniquely in the form

$$f(x) = \sum_{i=0}^{k} u_i b_i(x) = \sum_{i=0}^{k} u_i \binom{x}{i},$$

where $u_0, u_1, \ldots, u_k$ are integers and $u_k \neq 0$ (by Exercise 4). This is the standard representation of an integer-valued polynomial.

If $f(x)$ is an integer-valued polynomial of degree $k \geq 1$ with positive leading coefficient, then there exists a nonnegative integer m such that $f(m) \geq 0$ and $f(x)$ is strictly increasing for $x \geq m$. Let $f_m(x) = f(x+m)$. Then $f_m(x)$ is an integer-valued polynomial such that

$$A(f_m) = \{f_m(i)\}_{i=0}^{\infty}$$

is a strictly increasing sequence of nonnegative integers. The polynomials $f(x)$ and $f_m(x)$ have the same degrees and the same leading coefficients (by Exercise 9). Replacing $f(x)$ with $f_m(x)$, we can assume that $f(x)$ is an integer-valued polynomial such that

$$A(f) = \{f(i)\}_{i=0}^{\infty}$$

is a strictly increasing sequence of nonnegative integers.

Waring's problem for polynomials states that if the greatest common divisor of the set $A(f)$ is 1, then every sufficiently large integer can be written as the sum of a bounded number of elements of $A(f)$. If also $0, 1 \in A(f)$, then there exists an integer h such that every nonnegative integer can be written as the sum of exactly h elements of $A(f)$. The classical Waring's problem is the special case $f(x) = x^k$. We shall also prove Waring's problem for polynomials by Linnik's method.

In the next chapter we obtain a generalization of Waring's problem for finite sequences of polynomials.

Exercises

In this set of exercises we characterize integer-valued polynomials.

1. Define $b_0(x) = 1$. For every integer $k \geq 1$, define the kth *binomial polynomial*

$$b_k(x) = \binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}.$$

Compute $b_k(x)$ for $k = 0, 1, 2, 3$. Prove that if $k \geq 1$ and $n \geq 1$, then

$$b_k(-n) = (-1)^k b_k(n+k-1).$$

Prove that if $f(x)$ is a polynomial of degree k with complex coefficients, then there exist unique complex numbers u_0, $u_1, \ldots,$ u_k with $u_k \neq 0$ such that

$$f(x) = \sum_{i=0}^{k} u_i b_i(x) = \sum_{i=0}^{k} u_i \binom{x}{i}. \tag{11.1}$$

2. For any function $f(x)$, define the *difference operator*

$$\Delta f(x) = f(x+1) - f(x).$$

Prove that $\Delta b_0(x) = 0$ and that

$$\Delta b_k(x) = b_{k-1}(x)$$

for all $k \geq 1$. If

$$f(x) = \sum_{i=0}^{k} u_i \binom{x}{i},$$

prove that

$$\Delta f(x) = \sum_{i=0}^{k-1} u_{i+1} \binom{x}{i}.$$

3. A polynomial $f(x)$ is called *integer-valued* if $f(n)$ is an integer for every integer n, that is, if $f(\mathbf{Z}) \subseteq \mathbf{Z}$. Prove that $b_k(x)$ is an integer-valued polynomial of degree k for every $k \geq 0$. Prove that if u_0, $u_1, \ldots,$ u_k are integers and $u_k \neq 0$, then

$$f(x) = \sum_{i=0}^{k} u_i \binom{x}{i}$$

is an integer-valued polynomial of degree k.

4. Let $f(x)$ be a polynomial of degree k with complex coefficients. Prove that if $f(x)$ is an integer for all sufficiently large integers x, then there exist unique integers $u_0, u_1, \ldots, u_k$ with $u_k \neq 0$ such that

$$f(x) = \sum_{i=0}^{k} u_i \binom{x}{i}.$$

Hint: Observe that if $k \geq 1$ and $f(x)$ is integer-valued for all sufficiently large x, then $\Delta f(x)$ is also integer-valued for all sufficiently large x. Represent $f(x)$ in the form (11.1) and use induction on k.

5. Let $f(x)$ be a polynomial of degree k with complex coefficients. Prove that if $f(x)$ is an integer for all sufficiently large integers x, then $f(x)$ is an integer for all integers x.

6. Prove that if $f(x)$ is an integer-valued polynomial of degree k with leading coefficient a_k, then
$$|a_k| \geq \frac{1}{k!}.$$

7. Let $f(x)$ be an integer-valued polynomial, and define
$$d = \gcd\{f(x) : x \in \mathbf{N}_0\}$$
and
$$d' = \gcd\{f(x) : x \in \mathbf{Z}\}.$$
Let $u_0, u_1, \ldots, u_k$ be integers such that
$$f(x) = \sum_{i=0}^{k} u_i \binom{x}{i}.$$
Prove that
$$d = d' = (u_0, u_1, \ldots, u_k).$$

8. Prove that if
$$f(x) = \sum_{i=0}^{k} u_i \binom{x}{i},$$
then
$$f_1(x) = f(x+1) = u_k \binom{x}{k} + \sum_{i=0}^{k-1} (u_i + u_{i+1}) \binom{x}{i}.$$
Prove that
$$\begin{aligned} &\gcd(u_0, u_1, \ldots, u_{k-1}, u_k) \\ = \quad &\gcd(u_0 + u_1, u_1 + u_2, \ldots, u_{k-1} + u_k, u_k). \end{aligned}$$

9. Let $f(x)$ be an integer-valued polynomial and let $m \in \mathbf{Z}$. We define the polynomial $f_m(x) = f(x+m)$. Prove that $f(x)$ and $f_m(x)$ are polynomials of the same degree and with the same leading coefficient. Let $A(f) = \{f(i)\}_{i=0}^{\infty}$. Prove that $\gcd(A(f)) = \gcd(A(f_m))$.

11.2 Stable Bases

A set A of nonnegative integers is called a *basis of order* h if every positive integer can be written as the sum of exactly h elements of A. The set A is called a *basis of finite order* if A is a basis of order h for some h. For example, by Lagrange's theorem the set of squares is a basis of order four. Waring's problem states that for every $k \geq 2$, the set of nonnegative kth powers is a basis of finite order.

Let $A = \{a_i\}_{i=0}^{\infty}$ be an infinite set of nonnegative integers such that $a_0 < a_1 < a_2 \cdots$. The *counting function* of A, denoted by $A(n)$, counts the number of positive elements of A that do not exceed n, that is,

$$A(n) = \sum_{\substack{a_i \in A \\ 1 \leq a_i \leq n}} 1.$$

The *Shnirel'man density* of the set A is

$$\begin{aligned} \sigma(A) &= \inf\left\{ \frac{A(n)}{n} : n = 1, 2, \ldots \right\} \\ &= \sup\left\{ \alpha : \frac{A(n)}{n} \geq \alpha \text{ for all } n = 1, 2, \ldots \right\}. \end{aligned}$$

Then $0 \leq \sigma(A) \leq 1$ for every set A. If $\sigma(A) = \alpha$, then $A(n) \geq \alpha n$ for every $n \geq 1$.

Let $B = \{b_i\}_{i=0}^{\infty}$ be a set of nonnegative integers such that $0 = b_0 < b_1 < b_2 < \cdots$. We construct the subset $A_B \subseteq A$ as follows:

$$A_B = \{a_{b_i}\}_{i=0}^{\infty}.$$

Then

$$a_0 = a_{b_0} < a_{b_1} < a_{b_2} < \cdots.$$

For example, $A_{\mathbf{N}_0} = A$.

If the Shnirel'man density of B is positive, then A_B is called a *subset of A of positive Shnirel'man density.* The set A is called a *stable basis* if every subset of A of positive Shnirel'man density is a basis of finite order. Shnirel'man proved that the set of kth powers is a stable basis for every $k \geq 1$. We shall also prove this generalization of Waring's problem.

A set A of nonnegative integers is called an *asymptotic basis of order* h if every sufficiently large positive integer can be written as the sum of exactly h elements of A. We call A an *asymptotic basis of finite order* if A is an asymptotic basis of order h for some h. Let $\gcd(A)$ denote the greatest common divisor of the elements of the set A. If $\gcd(A) = d$, then every sum of elements of A is divisible by d. It follows that the set A is an asymptotic basis only if $\gcd(A) = 1$.

The *lower asymptotic density* of the set A is

$$d_L(A) = \liminf \left\{ \frac{A(n)}{n} : n = 1, 2, \ldots \right\}.$$

Then $0 \leq d_L(A) \leq 1$ for every set A. Let $B = \{b_i\}_{i=0}^{\infty}$ be a strictly increasing sequence of nonnegative integers. If the lower asymptotic density of B is positive, then the set A_B is called a *subset of A of positive lower asymptotic density.* An *asymptotically stable basis* is a set A that satisfies the following condition: If $d_L(B) > 0$ and $\gcd(A_B) = d$, then there exists an integer $h = h(B)$ such that every sufficiently large multiple of d can be written as the sum of at most h elements of A_B. In particular, A_B is an asymptotic basis of finite order for every set B such that $d_L(B) > 0$ and $\gcd(A_B) = 1$.

We shall also prove that the set of kth powers is an asymptotically stable basis for every $k \geq 1$.

Exercises

1. Let A be a set of nonnegative integers. Prove that if $\sigma(A) > 0$, then $1 \in A$.

2. Let $m \geq 2$. Let A_r be the set of all nonnegative integers a such that $a \equiv r \pmod{m}$. Compute the Shnirel'man density of A_r and the lower asymptotic density of A_r for $r = 0, 1, \ldots, m-1$.

3. For $k \geq 2$, let $A^{(k)} = \{n^k : n \in \mathbf{N}_0\}$ be the set of the kth powers of the nonnegative integers. Compute the Shnirel'man density of $A^{(k)}$.

4. Let $A^{(\infty)} = \cup_{k=2}^{\infty} A^{(k)}$, where $A^{(k)}$ is the set of kth powers. Compute the Shnirel'man density of $A^{(\infty)}$.

5. Let $\mathbf{P}$ be the set of prime numbers and let $\mathbf{P}' = \mathbf{P} \cup \{1\}$. Compute the Shnirel'man density of $\mathbf{P}'$.

6. Recall that $[x]$ denotes the integer part of the real number x. Let $L_0 = \{[\log n] : n = 1, 2, 3, \ldots\}$. Compute the Shnirel'man density of L_0.

7. Compute the Shnirel'man density of the set $L_1 = \{[n \log n] : n = 1, 2, 3, \ldots\}$.

8. For $0 < a < 1$, let $L_a = \{[n^a \log n] : n = 1, 2, 3, \ldots\}$. Compute the Shnirel'man density of the set L_a.

9. Let $A = \{a_i\}_{i=1}^{\infty}$ be a set of positive integers with $1 = a_1 < a_2 < a_3 < \cdots$. Prove that $\sigma(A) > 0$ if $\limsup_{i \to \infty}(a_{i+1} - a_i) < \infty$.

10. Let $A = \{a_i\}_{i=1}^{\infty}$ be a set of positive integers with $1 = a_1 < a_2 < a_3 < \cdots$. Prove that $\sigma(A) = 0$ if $\lim_{i\to\infty}(a_{i+1} - a_i) = \infty$.

11. Construct a set $A = \{a_i\}_{i=0}^{\infty}$ of positive integers such that $\sigma(A) > 0$ and $\limsup_{i\to\infty}(a_{i+1} - a_i) = \infty$.

12. Let $A = \{a_i\}_{i=0}^{\infty}$ and $B = \{b_i\}_{i=0}^{\infty}$ be infinite sets of nonnegative integers with

$$\begin{aligned} 0 &= a_0 < a_1 < a_2 < \cdots, \\ 0 &= b_0 < b_1 < b_2 < \cdots, \end{aligned}$$

and counting functions $A(n)$ and $B(n)$, respectively. Let $A_B(n)$ be the counting function of the set $A_B = \{a_{b_i}\}_{i=0}^{\infty}$. Prove that

$$\begin{aligned} A_B(n) &= B(A(n)), \\ \sigma(A_B) &\geq \sigma(A)\sigma(B), \end{aligned}$$

and

$$d_L(A_B) \geq d_L(A)d_L(B).$$

11.3 Shnirel'man's Theorem

Let A and B be nonempty sets of integers. The *sumset* $A + B$ is the set consisting of all integers of the form $a + b$, where $a \in A$ and $b \in B$. The *difference set* $A - B$ consists of all integers of the form $a - b$, where $a \in A$ and $b \in B$.

If $A_1, A_2, \ldots, A_h$ are h sets of integers, then

$$A_1 + A_2 + \cdots + A_h$$

denotes the sumset consisting of all integers of the form $a_1 + a_2 + \cdots + a_h$, where $a_i \in A_i$ for $i = 1, 2, \ldots, h$. If $A_i = A$ for all $i = 1, 2, \ldots, h$, we let

$$hA = \underbrace{A + \cdots + A}_{h \text{ times}}.$$

Then A is a basis of order h if $\mathbf{N}_0 \subseteq hA$, that is, if the sumset hA contains every nonnegative integer. The set A is an asymptotic basis of order h if hA contains every sufficiently large integer.

Let A be a set of integers. If A contains every positive integer, then $A(n) = n$ for all $n \geq 1$ and A has Shnirel'man density $\sigma(A) = 1$. If $n \notin A$ for some $n \geq 1$, then $A(n) \leq n - 1$ and

$$\sigma(A) \leq \frac{A(n)}{n} \leq 1 - \frac{1}{n} < 1.$$

Thus, $\sigma(A) = 1$ if and only if A contains every positive integer.

Shnirel'man density is an important additive measure of the size of a set of integers. In particular, the set A is a basis of order h if and only if $\sigma(hA) = 1$, and the set A is a basis of finite order if and only if $\sigma(hA) = 1$ for some $h \geq 1$. Shnirel'man made the simple but extraordinarily powerful discovery that if A is any set of integers that contains 0 and has positive Shnirel'man density, then A is a basis of finite order. It follows that if $\sigma(A) = 0$ but $\sigma(h_1 A) > 0$ for some integer h_1, then the sumset $h_1 A$ is a basis of finite order, and so A is also a basis of finite order. This is a key idea in our proof of Waring's problem. Although the set $A^{(k)}$ of nonnegative kth powers has Shnirel'man density zero, we shall prove that there exists an integer h_1 such that the set $h_1 A^{(k)}$ of all sums of h_1 nonnegative kth powers has positive Shnirel'man density.

Lemma 11.1 *Let A and B be sets of integers such that $0 \in A$ and $0 \in B$. If $A(n) + B(n) \geq n$, then $n \in A + B$.*

Proof. If $n \in A$, then $n = n + 0 \in A + B$. Similarly, if $n \in B$, then $n = 0 + n \in A + B$.

Suppose that $n \notin A \cup B$. Define sets A' and B' by

$$A' = \{n - a : a \in A, 1 \leq a \leq n - 1\}$$

and

$$B' = B \cap [1, n - 1].$$

Then $|A'| = A(n)$, since $n \notin A$, and $|B'| = B(n)$, since $n \notin B$. Moreover,

$$A' \cup B' \subseteq [1, n - 1].$$

Since

$$|A'| + |B'| = A(n) + B(n) \geq n,$$

it follows that

$$A' \cap B' \neq \emptyset.$$

Therefore, $n - a = b$ for some $a \in A$ and $b \in B$, and so $n = a + b \in A + B$. □

Lemma 11.2 *Let A and B be sets of integers such that $0 \in A$ and $0 \in B$. If $\sigma(A) + \sigma(B) \geq 1$, then $\mathbf{N}_0 \subseteq A + B$.*

Proof. We have $0 = 0 + 0 \in A + B$. If $n \geq 1$, then

$$A(n) + B(n) \geq (\sigma(A) + \sigma(B))n \geq n,$$

and Lemma 11.1 implies that $n \in A + B$. □

Lemma 11.3 *Let A be a set of integers such that $0 \in A$ and $\sigma(A) \geq 1/2$. Then A is a basis of order 2.*

Proof. This follows immediately from Lemma 11.2 with $A = B$. □

Theorem 11.1 (Shnirel'man) *Let A and B be sets of integers such that $0 \in A$ and $0 \in B$. Let $\sigma(A) = \alpha$ and $\sigma(B) = \beta$. Then*

$$\sigma(A + B) \geq \alpha + \beta - \alpha\beta. \tag{11.2}$$

Proof. Let $n \geq 1$. Let $a_0 = 0$ and let

$$1 \leq a_1 < \cdots < a_k \leq n$$

be the $k = A(n)$ positive elements of A that do not exceed n. Since $0 \in B$, it follows that $a_i = a_i + 0 \in A + B$ for $i = 1, \ldots, k$. For $i = 0, \ldots, k-1$, let

$$1 \leq b_1 < \cdots < b_{r_i} \leq a_{i+1} - a_i - 1$$

be the $r_i = B(a_{i+1} - a_i - 1)$ positive integers in B that are less than $a_{i+1} - a_i$. Then

$$a_i < a_i + b_1 < \cdots < a_i + b_{r_i} < a_{i+1}$$

and

$$a_i + b_j \in A + B$$

for $j = 1, \ldots, r_i$. Let

$$1 \leq b_1 < \cdots < b_{r_k} \leq n - a_k$$

be the $r_k = B(n - a_k)$ positive integers in B that do not exceed $n - a_k$. Then

$$a_k < a_k + b_1 < \cdots < a_k + b_{r_k} \leq n$$

and

$$a_k + b_j \in A + B$$

for $j = 1, \ldots, r_k$. It follows that

$$\begin{aligned} (A+B)(n) &\geq A(n) + \sum_{i=0}^{k} r_i \\ &= A(n) + \sum_{i=0}^{k-1} B(a_{i+1} - a_i - 1) + B(n - a_k) \\ &\geq A(n) + \beta \sum_{i=0}^{k-1} (a_{i+1} - a_i - 1) + \beta(n - a_k) \end{aligned}$$

$$\begin{aligned}
&= A(n) + \beta n - \beta k \\
&= (1-\beta)A(n) + \beta n \\
&\geq (1-\beta)\alpha n + \beta n \\
&= (\alpha + \beta - \alpha\beta)n,
\end{aligned}$$

and so

$$\frac{(A+B)(n)}{n} \geq \alpha \ + \beta \ - \alpha\beta$$

for all positive integers n. Therefore,

$$\sigma(A+B) = \inf\left\{\frac{(A+B)(n)}{n} : n = 1, 2, \ldots\right\} \geq \alpha + \beta - \alpha\beta.$$

This completes the proof. □

Inequality (11.2) can be expressed as follows:

$$1 - \sigma(A+B) \leq (1-\sigma(A))(1-\sigma(B)). \tag{11.3}$$

We can generalize this inequality to the sum of any finite number of sets of integers.

Theorem 11.2 *Let $h \geq 1$, and let $A_1, \ldots, A_h$ be sets of integers with $0 \in A_i$ for $i = 1, \ldots, h$. Then*

$$1 - \sigma(A_1 + \cdots + A_h) \leq \prod_{i=1}^{h}(1 - \sigma(A_i)).$$

Proof. This is by induction on h. Let $\sigma(A_i) = \alpha_i$ for $i = 1, \ldots, h$. For $h = 1$, there is nothing to prove, and for $h = 2$ the inequality is equivalent to (11.3).

Let $h \geq 3$, and assume that the theorem holds for $h - 1$ sets. Let $A_1, \ldots, A_h$ be h sets of integers such that $0 \in A_i$ for all i. Let $B = A_1 + \cdots + A_{h-1}$. We have the induction hypothesis

$$1 - \sigma(B) = 1 - \sigma(A_1 + \cdots + A_{h-1}) \leq \prod_{i=1}^{h-1}(1 - \sigma(A_i)),$$

and so

$$\begin{aligned}
1 - \sigma(A_1 + \cdots + A_h) &= 1 - \sigma(B + A_h) \\
&\leq (1 - \sigma(B)(1 - \sigma(A_h)) \\
&\leq \prod_{i=1}^{h-1}(1 - \sigma(A_i))(1 - \sigma(A_h) \\
&= \prod_{i=1}^{h}(1 - \sigma(A_i)).
\end{aligned}$$

This completes the proof. □

Theorem 11.3 *Let* $0 < \alpha \leq 1$. *There exists an integer* $h = h(\alpha)$ *such that if* $A_1, \ldots, A_h$ *are sets of nonnegative integers with* $0 \in A_i$ *and* $\sigma(A_i) \geq \alpha$ *for all* $i = 1, \ldots, h$, *then*

$$A_1 + \cdots + A_h = \mathbf{N}_0.$$

Proof. Since $0 \leq 1 - \alpha < 1$, there exists a positive integer h_1 such that

$$0 \leq (1-\alpha)^{h_1} \leq \frac{1}{2}.$$

Let $h = 2h_1$, and let $A_1, \ldots, A_h$ be sets of nonnegative integers with $0 \in A_i$ and $\sigma(A_i) \geq \alpha$ for $i = 1, \ldots, h$. We define $A = A_1 + \cdots + A_{h_1}$ and $B = A_{h_1+1} + \cdots + A_{2h_1}$. By Theorem 11.2,

$$\sigma(A) = \sigma(A_1 + \cdots + A_{h_1}) \geq 1 - \prod_{i=1}^{h_1}(1 - \sigma(A_i)) \geq 1 - (1-\alpha)^{h_1} \geq \frac{1}{2}.$$

Similarly,

$$\sigma(B) = \sigma(A_{h_1+1} + \cdots + A_{2h_1}) \geq \frac{1}{2}.$$

Applying Lemma 11.3, we obtain

$$A_1 + \cdots + A_h = A + B = \mathbf{N}_0.$$

This completes the proof. □

Theorem 11.4 (Shnirel'man) *Let* A *be a set of nonnegative integers such that* $0 \in A$ *and* $\sigma(A) > 0$. *Then* A *is a basis of finite order.*

Proof. Let $\alpha = \sigma(A)$. The result follows from Theorem 11.3 with $A_i = A$ for $i = 1, \ldots, h(\alpha)$.

Theorem 11.5 *Let* A *be a set of nonnegative integers with* $0 \in A$ *such that* $\sigma(h_1 A) > 0$ *for some positive integer* h_1. *Then* A *is a basis of finite order.*

Proof. If $\sigma(h_1 A) > 0$, then there exists an integer h_2 such that $h_1 A$ is a basis of order h_2, that is, every nonnegative integer is a sum of h_2 elements of $h_1 A$. Since

$$h_2(h_1 A) = (h_1 h_2)A,$$

the set A is a basis of order $h = h_1 h_2$. □

Theorem 11.6 *Let B be a set of nonnegative integers with $0 \in B$ and $\gcd(B) = 1$. If $d_L(B) > 0$, then B is an asymptotic basis of finite order.*

Proof. The set $A = B \cup \{1\}$ has positive Shnirel'man density (by Exercise 1), and so A is a basis of order h_1 for some positive integer h_1. It follows that every nonnegative integer can be written in the form $u + j$, where $0 \leq j \leq h_1$ and u is a sum of $h_1 - j$ elements of B. Since $0 \in B$,

$$u \in (h_1 - j)B \subseteq h_1 B.$$

If B is any set of relatively prime positive integers, then, by Theorem 1.16, there exists an integer $n_0 = n_0(B)$ such that every integer $n \geq n_0$ can be represented as a sum of elements of B. Since $0 \in B$ and $\gcd(B) = 1$, there exists a positive integer h_2 such that

$$n_0 + j \in h_2 B$$

for $j = 0, 1, \ldots, h_1$. Let $h = h_1 + h_2$. If $n \geq n_0$, then $n - n_0 \geq 0$ and we can write $n - n_0$ in the form $u + j$, where $u \in h_1 B$ and $0 \leq j \leq h_1$. Then

$$n = u + (n_0 + j) \in h_1 B + h_2 B = hB,$$

and so B is an asymptotic basis of finite order. □

Theorem 11.7 *Let B be a set of nonnegative integers with $\gcd(B) = d$. If $d_L(B) > 0$, then every sufficiently large multiple of d is the sum of a bounded number of elements of B.*

Proof. The set $d^{-1} * B = \{b/d : b \in B\}$ consists of nonnegative integers, and

$$A = \{0\} \cup d^{-1} * B$$

is a set of nonnegative integers with $0 \in A$ and $\gcd(A) = 1$. By Theorem 11.6, every sufficiently large integer can be represented as the sum of exactly h elements of A, and so every sufficiently large multiple of d can be represented as the sum of at most h elements of B. □

Exercises

1. Let A be a set of nonnegative integers. Prove that $\sigma(A) > 0$ if and only if $1 \in A$ and $d_L(A) > 0$.

2. Let h_1 and h_2 be positive integers with $h_1 < h_2$, and let A be a nonempty set of integers. Prove that

$$h_1 A + h_2 A = (h_1 + h_2)A.$$

Prove that

$$h_1A - h_2A = (h_1 - h_2)A$$

if and only if $|A| = 1$.

3. Let A be a set of nonnegative integers such that $0 \in A$ and

$$0 < \sigma(A) \leq \frac{1}{2}.$$

Prove that

$$\sigma(2A) \geq \frac{3}{2}\sigma(A).$$

Use this to give another proof of Theorem 11.4.

4. Let A be a set of nonnegative integers such that $0 \in A$, $A \neq \{0\}$, and $hA = (h+1)A$ for some positive integer h.

   (a) Prove that $hA = \ell A$ for all $\ell \geq h$.

   (b) Prove that hA is periodic, that is, there exists a positive integer m such that if $b \in hA$, then $b + m \in hA$.

   (c) Let $d = \gcd(A)$. Prove that $hA \sim d * \mathbf{N}_0$, that is, the sumset hA eventually coincides with the set of all multiples of d.

11.4 Waring's Problem for Polynomials

Let $f(x)$ be an integer-valued polynomial of degree k such that

$$A(f) = \{f(i)\}_{i=0}^{\infty}$$

is a strictly increasing sequence of nonnegative integers. Let d be the greatest common divisor of $A(f)$. By Exercises 5 and 7 in Section 11.1, the polynomial $f(x)/d$ is also integer-valued of degree k, and the greatest common divisor of $A(f(x)/d)$ is 1. Without loss of generality, we can assume that $f(x)$ is an integer-valued polynomial with $\gcd(A(f)) = 1$.

Let NSE denote "*the number of solutions of the equation.*" We define *representation functions* $r_{f,s}(n)$ and $R_{f,s}(N)$ for the polynomial $f(x)$ by

$$r_{f,s}(n) = \text{NSE } \{f(x_1) + \cdots + f(x_s) = n : x_1, \ldots, x_s \in \mathbf{N}_0\}$$

and

$$R_{f,s}(N) = \sum_{0 \leq n \leq N} r_{f,s}(n).$$

Lemma 11.4 *Let* $f(x) = \sum_{i=0}^{k} a_i x^i$ *be an integer-valued polynomial of degree* k *with leading coefficient* $a_k > 0$. *Let*

$$x^*(f) = \frac{2(|a_{k-1}| + |a_{k-2}| + \cdots + |a_0|)}{a_k}. \tag{11.4}$$

If $x > x^*(f)$ *is an integer, then*

$$\frac{a_k x^k}{2} < f(x) < \frac{3a_k x^k}{2}. \tag{11.5}$$

If N *is sufficiently large, then*

$$R_{f,s}(N) > \frac{1}{2}\left(\frac{2N}{3a_k s}\right)^{s/k}. \tag{11.6}$$

Proof. Since

$$f(x) = a_k x^k \left(1 + \frac{a_{k-1}}{a_k x} + \frac{a_{k-2}}{a_k x^2} + \cdots + \frac{a_0}{a_k x^k}\right),$$

it follows for $x > x^*(f)$ that

$$\begin{aligned} \left|\frac{f(x)}{a_k x^k} - 1\right| &= \left|\frac{a_{k-1}}{a_k x} + \frac{a_{k-2}}{a_k x^2} + \cdots + \frac{a_0}{a_k x^k}\right| \\ &\leq \frac{|a_{k-1}|}{a_k x} + \frac{|a_{k-2}|}{a_k x^2} + \cdots + \frac{|a_0|}{a_k x^k} \\ &\leq \frac{|a_{k-1}| + |a_{k-2}| + \cdots + |a_0|}{a_k x} \\ &= \frac{x^*(f)}{2x} \\ &< \frac{1}{2}. \end{aligned}$$

This proves (11.5).

If $x_1, \ldots, x_s$ are integers such that

$$x^*(f) < x_j \leq \left(\frac{2N}{3a_k s}\right)^{1/k}$$

for $j = 1, \ldots, s$, then

$$0 < \frac{a_k x_j^k}{2} < f(x_j) < \frac{3a_k x_j^k}{2} \leq \frac{N}{s}$$

and

$$0 < f(x_1) + \cdots + f(x_s) < N.$$

The number of integers in the interval

$$\left(x^*(f), \left(\frac{2N}{3a_k s}\right)^{1/k}\right]$$

is greater than

$$\left(\frac{2N}{3a_k s}\right)^{1/k} - x^*(f) - 1,$$

and so

$$R_{f,s}(N) > \left(\left(\frac{2N}{3a_k s}\right)^{1/k} - x^*(f) - 1\right)^s \geq \frac{1}{2}\left(\frac{2N}{3a_k s}\right)^{s/k}$$

for N sufficiently large. This proves (11.6). □

Lemma 11.5 *Let $f(x) = \sum_{i=0}^k a_i x^i$ be an integer-valued polynomial of degree k such that*

$$A(f) = \{f(i)\}_{i=0}^\infty$$

is a strictly increasing sequence of nonnegative integers. Define $x^*(f)$ by (11.4) and let

$$N(f) = \frac{x^*(f)^k}{2k!}. \tag{11.7}$$

For $N \geq N(f)$, if $x_1, \ldots, x_s$ are nonnegative integers with

$$\sum_{j=1}^s f(x_j) \leq N,$$

then

$$0 \leq x_j \leq (2k!N)^{1/k} \qquad \textit{for } j = 1, \ldots, s.$$

Proof. Recall that $k!a_k \geq 1$ by Exercise 6 in Section 11.1. If $N \geq N(f)$ and $x_j > (2k!N)^{1/k} \geq x^*(f)$, then

$$f(x_j) > \frac{a_k x_j^k}{2} \geq k!a_k N \geq N,$$

and so

$$\sum_{i=1}^s f(x_i) \geq f(x_j) > N.$$

This completes the proof. □

A critical part of Linnik's solution of Waring's problem is the following result, which is a special case of Theorem 12.3.

Theorem 11.8 *Let $\{s(k)\}_{k=1}^{\infty}$ be the sequence of integers defined recursively by $s(1) = 1$ and*

$$s(k) = 8k2^{[\log_2 s(k-1)]} \qquad \text{for } k \geq 2.$$

Let $c \geq 1$ and $P \geq 1$. If

$$f(x) = \sum_{i=0}^{k} a_i x^i$$

is an integer-valued polynomial of degree k such that

$$|a_i| \leq cP^{k-i} \qquad \text{for } i = 0, 1, \ldots, k,$$

then for every integer n,

$$NSE \left\{ \begin{array}{l} \sum_{j=1}^{s(k)} f(x_j) = n \quad \text{with } x_j \in \mathbf{Z} \\ \text{and } |x_j| \leq cP \text{ for } j = 1, \ldots, s(k) \end{array} \right\} \ll_{k,c} P^{s(k)-k}.$$

Proof. Let $c = c_1$ and $f_j(x) = f(x)$ for $j = 1, \ldots, s(k)$ in Theorem 12.3. □

Theorem 11.9 *Let $f(x) = \sum_{i=0}^{k} a_i x^i$ be an integer-valued polynomial of degree k with $a_k > 0$ and $\gcd(A(f)) = 1$. Then $A(f) \cup \{0\}$ is an asymptotic basis of finite order, that is, for some h and every sufficiently large integer n there exists a positive integer $h_n \leq h$ and nonnegative integers $x_1, \ldots, x_{h_n}$ such that*

$$f(x_1) + \cdots + f(x_{h_n}) = n.$$

Proof. Define $N(f)$ by (11.7), and let $s = s(k)$ be the integer constructed in Theorem 11.8. Let $W = sA(f)$ be the set consisting of all sums of s integers of the form $f(x)$ with $x \in \mathbf{N}_0$. We shall prove that the sumset W has lower asymptotic density $d_L(W) > 0$.

Let $W(N)$ be the counting function of W. Choose $c \geq (2k!)^{1/k}$ and choose $N \geq N(f)$ sufficiently large that for $P = N^{1/k}$,

$$|a_i| \leq cP^{k-i} \qquad \text{for } i = 0, 1, \ldots, k.$$

Then $0 < a_k \leq c$. By Lemma 11.5, if $x_1, \ldots, x_s$ are nonnegative integers such that $\sum_{j=1}^{s} f(x_j) \leq N$, then

$$0 \leq x_j \leq (2k!N)^{1/k} \leq cP \qquad \text{for } j = 1, \ldots, s.$$

We get upper bounds for $r_{f,s}(n)$ and $R_{f,s}(N)$ as follows: If $0 \leq n \leq N$, then

$$\begin{aligned} r_{f,s}(n) &= \text{NSE}\ \{f(x_1) + \cdots + f(x_s) = n : x_i \in \mathbf{N}_0\} \\ &\leq \text{NSE}\ \{f(x_1) + \cdots + f(x_s) = n : |x_j| \leq cP\} \\ &\ll_{k,c} P^{s-k} \end{aligned}$$

by Theorem 11.8, and so

$$\begin{aligned} R_{f,s}(N) &= \sum_{0 \leq n \leq N} r_{f,s}(n) \\ &= \sum_{\substack{0 \leq n \leq N \\ r_{f,s}(n) \geq 1}} r_{f,s}(n) \\ &\ll_{k,c} W(N) P^{s-k} \\ &\ll_{k,c} \left(\frac{W(N)}{N} \right) P^s. \end{aligned}$$

We can apply Lemma 11.4 to obtain a lower bound for $R_{k,s}(N)$. For N sufficiently large,

$$R_{f,s}(N) > \frac{1}{2} \left(\frac{2N}{3a_k s} \right)^{s/k} \geq \frac{1}{2} \left(\frac{2N}{3cs} \right)^{s/k} \gg_{k,c} P^s.$$

Therefore,

$$P^s \ll_{k,c} R_{f,s}(N) \ll_{k,c} \left(\frac{W(N)}{N} \right) P^s,$$

and so $W(N)/N \gg_{k,c} 1$. It follows that

$$d_L(sA(f)) = d_L(W) > 0,$$

and the result follows immediately from Theorem 11.7. □

Theorem 11.10 *Let $f(x)$ be an integer-valued polynomial of degree k with leading coefficient $a_k > 0$. If $0, 1 \in A(f) = \{f(x) : x \in \mathbf{N}_0\}$, then $A(f)$ is a basis of finite order.*

Proof. This is a consequence of Theorem 11.9. □

Theorem 11.11 (Waring–Hilbert) *For every $k \geq 2$, the set of nonnegative kth powers is a basis of finite order.*

Proof. This is the special case of Theorem 11.10 applied to the polynomial $f(x) = x^k$. □

Theorem 11.12 *Let $f(x)$ be an integer-valued polynomial of degree k with leading coefficient $a_k > 0$ and $\gcd(A(f)) = 1$. Then $A(f) \cup \{0\}$ is an asymptotically stable asymptotic basis of finite order.*

Proof. This requires only minor modifications of the proof of Theorem 11.9. Let $A(f) = \{f(i)\}_{i=0}^{\infty}$, and let B be a set of nonnegative integers of lower asymptotic density $d_L(B) = \beta > 0$. Then

$$A_B = \{f(b) : b \in B\}.$$

Let $s = s(k)$ be the integer constructed in Theorem 11.8. The sumset $W_s = sA_B$ consists of all sums of s integers of the form $f(b)$ with $b \in B$. Let $W_s(N)$ be the counting function of the sumset W_s. Let $r_{f,s}^{(B)}(n)$ denote the number of solutions of the equation

$$f(b_1) + \cdots + f(b_s) = n$$

with $b_1, \ldots, b_s \in B$, and let

$$R_{f,s}^{(B)}(N) = \sum_{n=0}^{N} r_{f,s}^{(B)}(n).$$

We shall again compute upper and lower bounds for $R_{f,s}^{(B)}(n)$.

Choose real numbers $c \geq (2k!)^{1/k}$ and $N \geq N(f)$ such that for $P = N^{1/k}$,

$$|a_i| \leq cP^{k-i} \qquad \text{for } i = 1, \ldots, k.$$

By Theorem 11.8, we have the upper bound

$$\begin{aligned} R_{k,s}^{(B)}(N) &= \sum_{\substack{n=0 \\ r_{k,s}^{(B)}(n)\geq 1}}^{N} r_{k,s}^{(B)}(n) \leq \sum_{\substack{n=0 \\ r_{k,s}^{(B)}(n)\geq 1}}^{N} r_{k,s}(n) \\ &\ll_{k,c} W_B(N)P^{s-k} \\ &\ll_{k,c} \left(\frac{W_B(N)}{N}\right) P^s \end{aligned}$$

for all sufficiently large N.

To obtain a lower bound, we observe that the number of integers $b \in B$ such that

$$x^*(f) < b \leq \left(\frac{2N}{3a_k s}\right)^{1/k} \tag{11.8}$$

is

$$B\left(\left(\frac{2N}{3a_k s}\right)^{1/k}\right) - B(x^*(f)) \geq \left(\frac{\beta}{2}\right)\left(\frac{2N}{3a_k s}\right)^{1/k} - B(x^*(f)) \gg_{k,c} P$$

for sufficiently large N. By Lemma 11.4, if $b \in B$ satisfies inequality (11.8), then

$$0 \leq f(b) \leq \frac{N}{s},$$

and so

$$R_{f,s}^{(B)}(N) \gg_{k,c} P^s.$$

It follows that $W_B(N)/N \gg_{k,c} 1$, and so $W_B = sA_B$ has positive lower asymptotic density. The result now follows from Theorem 11.7. □

Theorem 11.13 *Let $f(x)$ be an integer-valued polynomial of degree k with leading coefficient $a_k > 0$. If $0, 1 \in A(f) = \{f(x) : x \in \mathbf{N}_0\}$, then $A(f)$ is a stable basis of finite order.*

Proof. This follows from Theorem 11.12. □

Theorem 11.14 (Waring–Shnirel'man) *For every $k \geq 2$, the set of nonnegative kth powers is a stable basis of finite order and an asymptotically stable asymptotic basis of finite order.*

Proof. This follows from Theorem 11.12. □

Exercises

1. Prove that every multiple of 6 can be written as the sum of a bounded number of integers of the form $x(x-1)(x-2)$ with $x \in \mathbf{N}_0$.
2. Prove that for every $k \geq 1$ there is an integer $h(k)$ such that every positive integer can be written as the sum of at most $h(k)$ kth powers of odd numbers.

11.5 Notes

Nathanson's *Additive Number Theory: The Classical Bases* [104] contains proofs of Lagrange's theorem that every number is the sum of four squares, and Wieferich's theorem that every number is the sum of nine cubes. A proof of Lagrange's theorem that depends on the geometry of numbers appears in Nathanson [103]. Jacobi's formula for the number of representations of an integer as the sum of four squares is Theorem 14.4 in Chapter 14 of this book.

In 1909 Hilbert [66] gave the first proof of Waring's problem for all exponents $k \geq 2$. Hardy and Littlewood [55, 56] developed a different method of proof and obtained an asymptotic formula for $r_{k,s}(n)$. Vinogradov [150] simplified and improved the circle method of Hardy and Littlewood, and

obtained new results on Waring's problem. Nathanson's book [104] gives Hilbert's proof of Waring's problem and also a proof of the Hardy–Littlewood asymptotic formula. Vaughan [148] is the standard reference on the circle method.

This chapter contains Linnik's elementary proof of Waring's problem. Linnik [93] published this proof in 1943. An exposition of Linnik's proof also appears in Khinchin [78]. Rieger [122] refined Linnik's method to obtain an upper bound for the smallest integer $g(k)$ such that every nonnegative integer is the sum of $g(k)$ kth powers. This upper bound is much larger than the upper bound obtained by the circle method.

Kamke [76] proved Waring's problem for polynomials. Nechaev [109] has applied classical analytic techniques, that is, exponential sums and the circle method, to Waring's problem for polynomials. Kuzel' [86] observed that Linnik's method for the classical Waring's problem also applies to Waring's problem for polynomials.

12

Sums of Sequences of Polynomials

12.1 Sums and Differences of Weighted Sets

In this chapter we complete our study of Waring's problem by Linnik's method. We shall derive a fundamental upper bound for the number of representations of an integer as a sum of polynomials. In Chapter 11 we applied a special case of this result to solve Waring's problem for a single polynomial. In Section 12.4 we shall use the full strength of this upper bound to obtain a generalization of Waring's problem to sequences of polynomials.

We begin with the definition of a *weighted set*. A weighted set is a pair (A, w_A), where A is a set and w_A is a function (called the *weight function*) defined on A. In this chapter weighted sets are always finite sets of integers, and the range of the weight functions is the set of nonnegative integers, that is, $w_A(a) \in \mathbf{N}_0$ for all $a \in A$. Thus, we can think of a weighted set as a set with multiplicities, that is, a set in which the element a occurs or is counted $w_A(a)$ times.

There are natural ways to generate weighted sets. If (A, w_A) is a weighted set and A is a subset of A^*, then we can define the weighted set (A^*, w_{A^*}) by

$$w_{A^*}(a) = \begin{cases} w_A(a) & \text{if } a \in A, \\ 0 & \text{if } a \in A^* \setminus A. \end{cases} \tag{12.1}$$

Let $(A_1, w_{A_1}), \ldots, (A_h, w_{A_h})$ be weighted sets. The product set $A_1 \times \cdots \times A_h$ consists of all $h$tuples $(a_1, \ldots, a_h)$ with $a_i \in A_i$ for $i = 1, \ldots, h$.

We define a weight function on the product set by

$$w_{A_1\times\cdots\times A_h}(a_1,\ldots,a_h) = w_{A_1}(a_1)\cdots w_{A_h}(a_h).$$

Let $f : A_1 \times \cdots \times A_h \to B$ be a function defined on the product set. We define a weight function $w_B^{(f)}$ on B as follows:

$$\begin{aligned} w_B^{(f)}(b) &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=b}} w_{A_1\times\cdots\times A_h}(a_1,\ldots,a_h) \\ &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=b}} w_{A_1}(a_1)\cdots w_{A_h}(a_h). \end{aligned}$$

We can think of $w_B^{(f)}(b)$ as counting the weighted number of solutions of the equation $f(a_1,\ldots,a_h) = b$.

For example, if $A_1,\ldots,A_h$ are weighted sets of integers, then the sumset

$$S = A_1 + \cdots + A_h$$

is the image of the function $\sigma(a_1,\ldots,a_h) = a_1 + \cdots + a_h$ defined on the weighted product set $A_1 \times \cdots \times A_h$. The weight of an element $s \in S$ is

$$w_S^{(\sigma)}(s) = \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ a_1+\cdots+a_h=s}} w_{A_1}(a_1)\cdots w_{A_h}(a_h).$$

If $w_{A_i}(a_i) = 1$ for all $i = 1,\ldots,h$ and $a_i \in A_i$, then $w_S^{(\sigma)}(s)$ is simply the number of representations of s in the form $a_1 + \cdots + a_h$. Similarly, if we define $\delta : A_1 \times A_2 \to A_1 - A_2$ by $\delta(a_1,a_2) = a_1 - a_2$, then the difference set

$$D = A_1 - A_2 = \{a_1 - a_2 : a_1 \in A_1, a_2 \in A_2\}$$

is a weighted set of integers such that the weight of $d \in D$ is

$$w_D^{(\delta)}(d) = \sum_{\substack{(a_1,a_2)\in A_1\times A_2 \\ a_1-a_2=d}} w_{A_1}(a_1)w_{A_2}(a_2).$$

Let NSE denote "*the number of solutions of the equation*." If f is a function from the product set $A_1 \times \cdots \times A_h$ into a set B, then

$$\text{NSE}\left\{ \begin{array}{l} f(a_1,\ldots,a_h) = b \\ \text{with } a_i \in A_i \text{ for } i = 1,\ldots,h \end{array} \right\} = \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=b}} 1.$$

If $(A_1, w_{A_1}),\ldots,(A_h, w_{A_h})$ are weighted sets with $w_{A_i}(a_i) = 1$ for all $i = 1,\ldots,h$ and $a_i \in A_i$, then

$$w_B^{(f)}(b) = \text{NSE}\left\{ \begin{array}{l} f(a_1,\ldots,a_h) = b \\ \text{with } a_i \in A_i \text{ for } i = 1,\ldots,h \end{array} \right\}.$$

If w_i^* is an upper bound for the weight function w_{A_i}, that is, if $w_{A_i}(a_i) \leq w_i^*$ for all $i = 1, \ldots, h$ and $a_i \in A_i$, then

$$\begin{aligned} w_B^{(f)}(b) &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=b}} w_{A_1}(a_1)\cdots w_{A_h}(a_h) \\ &\leq \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=b}} w_1^*\cdots w_h^* \\ &= w_1^*\cdots w_h^* \text{ NSE}\left\{ \begin{array}{l} f(a_1,\ldots,a_h) = b \\ \text{with } a_i \in A_i \text{ for } i = 1,\ldots,h \end{array} \right\}. \end{aligned}$$

For brevity, we shall often refer to the weighted set (A, w_A) as the weighted set A.

Let A_1, A_2, and A_3 be weighted sets. We can form the weighted sumsets $S_1 = A_1 + A_2$ and $S_2 = A_2 + A_3$, and from these the weighted sumsets $S_1 + A_3$ and $A_1 + S_2$. We also have the weighted sumset $S = A_1 + A_2 + A_3$. By the associativity of set addition we have $S = S_1 + A_3 = A_1 + S_2$ as *sets*. In fact, these sets are also equal as *weighted sets*, that is, for every $s \in S$ we have

$$w_S(s) = w_{S_1+A_3}(s) = w_{A_1+S_2}(s). \tag{12.2}$$

This is a special case of the following theorem, which shows that weights constructed by composition of functions are well-defined.

Theorem 12.1 *For $\ell \geq 2$, let $h, r_0, r_1 \ldots, r_\ell$ be integers such that*

$$0 = r_0 < r_1 < \cdots < r_\ell = h.$$

Let $(A_1, w_{A_1}), \ldots, (A_h, w_{A_h})$ be weighted sets and let $B_1, \ldots, B_\ell$, and C be sets. For $i = 1, \ldots, \ell$, let

$$f_i : A_{r_{i-1}+1} \times \cdots \times A_{r_i} \to B_i$$

be a function defined on the weighted product set $A_{r_{i-1}+1} \times \cdots \times A_{r_i}$. Then f_i induces a weight function $w_{B_i}^{(f_i)}$ on the set B_i, and these weight functions determine a weight function on the product set $B_1 \times \cdots \times B_\ell$. Let

$$g : B_1 \times \cdots \times B_\ell \to C$$

be a function defined on the weighted product set $B_1 \times \cdots \times B_\ell$. Then g induces a weight function $w_C^{(g)}$ on C. Define the function

$$f : A_1 \times \cdots \times A_h \to C$$

by

$$\begin{aligned} &f(a_1, \ldots, a_h) \\ &\quad = g(f_1(a_1, \ldots, a_{r_1}), f_2(a_{r_1+1}, \ldots, a_{r_2}), \ldots, f_\ell(a_{r_{\ell-1}+1}, \ldots, a_{r_\ell}). \end{aligned}$$

Then f induces a weight function $w_C^{(f)}$ on C. For all $c \in C$ we have

$$w_C^{(f)}(c) = w_C^{(g)}(c),$$

that is,

$$\begin{aligned} &\sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=c}} w_{A_1\times\cdots\times A_h}(a_1,\ldots,a_h) \\ &\quad = \sum_{\substack{(b_1,\ldots,b_\ell)\in B_1\times\cdots\times B_\ell \\ g(b_1,\ldots,b_\ell)=c}} w_{B_1\times\cdots\times B_\ell}(b_1,\ldots,b_\ell). \end{aligned}$$

Proof. This is a straightforward calculation. We have

$$\begin{aligned} w_C^{(g)}(c) &= \sum_{\substack{(b_1,\ldots,b_\ell)\in B_1\times\cdots\times B_\ell \\ g(b_1,\ldots,b_\ell)=c}} w_{B_1\times\cdots\times B_\ell}(b_1,\cdots,b_\ell) \\ &= \sum_{\substack{(b_1,\ldots,b_\ell)\in B_1\times\cdots\times B_\ell \\ g(b_1,\ldots,b_\ell)=c}} w_{B_1}^{(f_1)}(b_1)\cdots w_{B_\ell}^{(f_\ell)}(b_\ell) \\ &= \sum_{\substack{(b_1,\ldots,b_\ell)\in B_1\times\cdots\times B_\ell \\ g(b_1,\ldots,b_\ell)=c}} \left(\sum_{\substack{(a_1,\ldots,a_{r_1})\in A_1\times\cdots\times A_{r_1} \\ f_1(a_1,\ldots,a_{r_1})=b_1}} \prod_{i=1}^{r_1} w_{A_i}(a_i) \right) \times \cdots \\ &\qquad \times \left(\sum_{\substack{(a_{r_{\ell-1}+1},\ldots,a_{r_\ell})\in A_{r_{\ell-1}+1}\times\cdots\times A_{r_\ell} \\ f_\ell(a_{r_{\ell-1}+1},\ldots,a_{r_\ell})=b_\ell}} \prod_{i=r_{\ell-1}+1}^{r_\ell} w_{A_i}(a_i) \right) \\ &= \sum_{\substack{(b_1,\ldots,b_\ell)\in B_1\times\cdots\times B_\ell \\ g(b_1,\ldots,b_\ell)=c}} \sum_{\substack{(a_1,\ldots,a_{r_1})\in A_1\times\cdots\times A_{r_1} \\ f_1(a_1,\ldots,a_{r_1})=b_1}} \cdots \\ &\qquad \sum_{\substack{(a_{r_{\ell-1}+1},\ldots,a_{r_\ell})\in A_{r_{\ell-1}+1}\times\cdots\times A_{r_\ell} \\ f_\ell(a_{r_{\ell-1}+1},\ldots,a_{r_\ell})=b_\ell}} \prod_{i=1}^{h} w_{A_i}(a_i) \\ &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ g\left(f_1(a_1,\ldots,a_{r_1}),\ldots,f_\ell(a_{r_{\ell-1}+1},\ldots,a_{r_\ell})\right)=c}} \prod_{i=1}^{h} w_{A_i}(a_i) \\ &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=c}} \prod_{i=1}^{h} w_{A_i}(a_i) \\ &= \sum_{\substack{(a_1,\ldots,a_h)\in A_1\times\cdots\times A_h \\ f(a_1,\ldots,a_h)=c}} w_{A_1\times\cdots\times A_h}(a_1,\ldots,a_h) \end{aligned}$$

$$= \quad w_C^{(f)}(c).$$

This completes the proof. □

Lemma 12.1 *Let B_1 and B_2 be weighted sets of integers. Define the addition map $\sigma : B_1 \times B_2 \to B_1 + B_2$ by $\sigma(b_1, b_2) = b_1 + b_2$ and the difference maps $\delta_i : B_i \times B_i \to B_i - B_i$ by $\delta_i(b_i, b_i') = b_i - b_i'$ for $i = 1, 2$. Consider the weighted sumset $S = B_1 + B_2$ and the weighted difference sets $D_1 = B_1 - B_1$ and $D_2 = B_2 - B_2$. Then for all integers n,*

$$w_S^{(\sigma)}(n) \leq \frac{1}{2}\left(w_{D_1}^{(\delta_1)}(0) + w_{D_2}^{(\delta_2)}(0)\right).$$

Proof. For $i = 1, 2$ we have

$$w_{D_i}^{(\delta_i)}(0) = \sum_{\substack{(b_i, b_i') \in B_i \times B_i \\ b_i - b_i' = 0}} w_{B_i}(b_i) w_{B_i}(b_i') = \sum_{b_i \in B_i} w_{B_i}(b_i)^2.$$

To each $b_1 \in B_1$ there exists at most one $b_2 \in B_2$ such that $b_1 + b_2 = n$. Applying the elementary inequality

$$xy \leq \frac{1}{2}\left(x^2 + y^2\right) \qquad \text{for } x, y \in \mathbf{R},$$

we obtain

$$\begin{aligned}
w_S^{(\sigma)}(n) &= \sum_{\substack{(b_1, b_2) \in B_1 \times B_2 \\ b_1 + b_2 = n}} w_{B_1}(b_1) w_{B_2}(b_2) \\
&\leq \sum_{\substack{(b_1, b_2) \in B_1 \times B_2 \\ b_1 + b_2 = n}} \frac{1}{2}\left(w_{B_1}(b_1)^2 + w_{B_2}(b_2)^2\right) \\
&\leq \frac{1}{2}\left(\sum_{b_1 \in B_1} w_{B_1}(b_1)^2 + \sum_{b_2 \in B_2} w_{B_2}(b_2)^2\right) \\
&= \frac{1}{2}\left(w_{D_1}^{(\delta_1)}(0) + w_{D_2}^{(\delta_2)}(0)\right).
\end{aligned}$$

This completes the proof. □

Lemma 12.2 *For $t \geq 1$, let $B_1, \ldots, B_{2^t}$ be weighted sets of integers, and let S be the weighted sumset*

$$S = B_1 + \cdots + B_{2^t}$$

with weight function determined by the addition map $\sigma : B_1 \times \cdots \times B_{2^t} \to B_1 + \cdots + B_{2^t}$. *For* $i = 1, \ldots, 2^t$, *consider the weighted difference sets*

$$D_i = 2^{t-1}B_i - 2^{t-1}B_i = 2^{t-1}(B_i - B_i)$$

with weight functions defined by the maps

$$\delta_i : B_i \times \cdots \times B_i \to D_i,$$

$$\delta_i(b_{i,1}, \ldots, b_{i,2^t}) = (b_{i,1} + \cdots + b_{i,2^{t-1}}) - (b_{i,2^{t-1}+1} + \cdots + b_{i,2^t}).$$

Then for all integers n,

$$w_S^{(\sigma)}(n) \leq \frac{1}{2^t} \sum_{i=1}^{2^t} w_{D_i}^{(\delta_i)}(0). \tag{12.3}$$

Let B *be a weighted set with weighted sumset* $S = 2^t B$ *and weighted difference set* $D = 2^{t-1}B - 2^{t-1}B$. *Then*

$$w_S^{(\sigma)}(n) \leq w_D^{(\delta)}(0) \tag{12.4}$$

for all integers $n \in S$.

Proof. The proof of (12.3) is by induction on t. The case $t = 1$ is Lemma 12.1.

Let $t \geq 2$, and assume that the lemma holds for $t - 1$. Consider the weighted sumsets

$$S_1 = B_1 + \cdots + B_{2^{t-1}}$$

and

$$S_2 = B_{2^{t-1}+1} + \cdots + B_{2^t}$$

with weights $w_{S_1}^{(\sigma_1)}$ and $w_{S_2}^{(\sigma_2)}$, respectively, and the weighted difference sets

$$T_1 = S_1 - S_1$$

and

$$T_2 = S_2 - S_2$$

with weights $w_{T_1}^{(\Delta_1)}$ and $w_{T_2}^{(\Delta_2)}$, respectively. Since

$$S = S_1 + S_2,$$

we can define an addition map $\sigma' : S_1 \times S_2 \to S$. By Theorem 12.1,

$$w_S^{(\sigma)}(s) = w_S^{(\sigma')}(s)$$

for all $s \in S$. (Indeed, Theorem 12.1 implies that all of the weight functions constructed in this proof are well-defined.)

By Lemma 12.1,

$$w_S^{(\sigma)}(s) \leq \frac{1}{2}\left(w_{T_1}^{(\Delta_1)}(0) + w_{T_2}^{(\Delta_2)}(0)\right)$$

for all $s \in S$. For $i = 1, \ldots, 2^t$, we define the weighted difference sets

$$B_i' = B_i - B_i.$$

Then

$$\begin{aligned} T_1 &= S_1 - S_1 \\ &= (B_1 + \cdots + B_{2^{t-1}}) - (B_1 + \cdots + B_{2^{t-1}}) \\ &= (B_1 - B_1) + \cdots + (B_{2^{t-1}} - B_{2^{t-1}}) \\ &= B_1' + \cdots + B_{2^{t-1}}'. \end{aligned}$$

Similarly,

$$T_2 = S_2 - S_2 = B_{2^{t-1}+1}' + \cdots + B_{2^t}'.$$

For $i = 1, \ldots, 2^t$, we define the weighted difference sets

$$D_i' = 2^{t-2}B_i' - 2^{t-2}B_i'$$

with weight functions $w_{D_i'}^{(\delta_i')}$. By induction, the lemma holds for sums of 2^{t-1} weighted sets. Therefore, we have

$$w_{T_1}^{(\Delta_1)}(0) \leq \frac{1}{2^{t-1}} \sum_{i=1}^{2^{t-1}} w_{D_i'}^{(\delta_i')}(0)$$

and

$$w_{T_2}^{(\Delta_2)}(0) \leq \frac{1}{2^{t-1}} \sum_{i=2^{t-1}+1}^{2^t} w_{D_i'}^{(\delta_i')}(0),$$

and so

$$w_S^{(\sigma)}(n) \leq \frac{1}{2}(w_{T_1}^{(\Delta_1)}(0) + w_{T_2}^{(\Delta_2)}(0)) = \frac{1}{2^t} \sum_{i=1}^{2^t} w_{D_i'}^{(\delta_i')}(0).$$

Since

$$\begin{aligned} D_i' &= 2^{t-2}B_i' - 2^{t-2}B_i' \\ &= 2^{t-2}(B_i - B_i) - 2^{t-2}(B_i - B_i) \\ &= 2^{t-1}B_i - 2^{t-1}B_i \\ &= D_i, \end{aligned}$$

it follows that

$$w_S^{(\sigma)}(n) \leq \frac{1}{2^t} \sum_{i=1}^{2^t} w_{D_i}^{(\delta_i)}(0).$$

Inequality (12.4) follows immediately from (12.3). □

Exercises

1. Let $A = \{0, 1, 3, 4\}$ be a weighted set with weight function $w_A(a) = 1$ for all $a \in A$. Compute the weight functions of the weighted sumset $2A$ and the weighted difference set $A - A$.

2. Let $A = \{0, 1, 3, 4\}$ be a weighted set with weight function $w_A(a) = a$ for all $a \in A$. Compute the weight functions of the weighted sumset $2A$ and the weighted difference set $A - A$.

3. Let $A = \{1, 2, 3, 4, 5\}$ be a weighted set with $w_A(a) = 1$ for all $a \in A$. Define $f : A \to A$ by $f(1) = f(2) = 3$ and $f(3) = f(4) = f(5) = 2$. Compute $w_A^{(f)}(a)$.

4. Let (A, w_A) be a weighted set, let $f : A \to B$ be a function, and let $w_B^{(f)}$ be the weight function induced on B by f. Prove that
$$\sum_{a\in A} w_A(a) = \sum_{b\in B} w_B^{(f)}(b).$$

5. Let $A = \{1, 2, 3, \ldots, n\}$ and let w_A be a weight function on A. Let S_n be the group of all permutations of A. If $\tau \in S_n$, then $\tau : A \to A$ induces a weight function $w_A^{(\tau)}$ on A. Prove that $w_A^{(\tau)}(a) = w_A(a)$ for all $\tau \in S_n$ and $a \in A$ if and only if w_A is a constant function.

6. Prove that Theorem 12.1 implies equation (12.2).

7. Let A be a weighted set. Prove the weighted set identity
$$(A - A) - (A - A) = 2A - 2A.$$

8. Let A be a set of integers of cardinality k. Prove that
$$|A + A| \leq \frac{k^2 + k}{2}$$
and
$$|A - A| \leq k^2 - k + 1.$$
For every positive integer k, construct a set A such that $|A| = k$, $|A + A| = (k^2 + k)/2$, and $|A - A| = k^2 - k + 1$.

12.2 Linear and Quadratic Equations

In this section we obtain upper bounds for certain linear and quadratic diophantine equations.

Lemma 12.3 *Let $Q \geq 1$. Let $u_1, \ldots, u_k$ be relatively prime integers such that*

$$U = \max\{|u_1|, \ldots, |u_k|\} \leq Q.$$

For every integer m,

$$NSE \left\{ \begin{array}{l} u_1v_1 + \cdots + u_kv_k = m \\ \text{with } |v_1|, \ldots, |v_k| \leq Q \end{array} \right\} \leq \frac{(k-1)!(3Q)^{k-1}}{U}. \tag{12.5}$$

Equivalently, for $i = 1, \ldots, k$ we can define the weighted sets $A_i = \{v \in \mathbf{Z} : |v| \leq Q\}$ with weights $w_{A_i}(v) = 1$ for all $v \in A_i$. Let B be the range of the function $f(v_1, \ldots, v_k) = u_1v_1 + \cdots + u_kv_k$. The lemma asserts that $w_B^{(f)}(m) \leq (k-1)!(3Q)^{k-1}/U$.

If we choose any $k-1$ numbers $v_1, \ldots, v_{k-1}$, then there exists at most one number v_k that satisfies the equation $u_1v_1 + \cdots + u_kv_k = m$. This gives the trivial upper bound $(2Q+1)^{k-1} \leq (3Q)^{k-1}$ for (12.5). A nontrivial assertion of the lemma is the denominator U in Q^{k-1}/U.

Proof. The proof is by induction on k. If $k = 1$, then $\gcd(u_1) = 1$ and $U = |u_1| = 1$. The number of solutions of the equation $u_1v_1 = m$ with $|v_1| \leq Q$ is at most

$$1 = \frac{0!(3Q)^0}{U}.$$

Let $k = 2$ and $U = \max\{|u_1|, |u_2|\} = |u_2|$. If

$$u_1v_1 + u_2v_2 = m, \tag{12.6}$$

then

$$u_1v_1 \equiv m \pmod{U}.$$

Since $(u_1, u_2) = (u_1, U) = 1$, we have

$$v_1 \equiv u_1^{-1}m \pmod{U}.$$

The number of integers v_1 in the congruence class $u_1^{-1}m \pmod{U}$ with $|v_1| \leq Q$ is at most

$$\frac{2Q}{U} + 1 \leq \frac{3Q}{U} \qquad \text{(since } U \leq Q\text{)}.$$

For each such integer v_1 there is at most one integer v_2 that satisfies the linear equation (12.6). Therefore,

$$\text{NSE } \{u_1v_1 + u_2v_2 = m \quad \text{with } |v_1|, |v_2| \leq Q\} \leq \frac{3Q}{U}.$$

Let $k \geq 3$, and assume that the lemma holds for $k-1$. Let

$$U = \max\{u_1, \ldots, u_k\} = |u_k|.$$

If $u_i = 0$ for $i = 1, \ldots, k-1$, then $1 = (u_1, \ldots, u_{k-1}, u_k) = |u_k| = U$, and the number of solutions of (12.5) is at most

$$(2Q+1)^{k-1} \leq (3Q)^{k-1} \leq \frac{(k-1)!(3Q)^{k-1}}{U}.$$

If $u_i \neq 0$ for some $i \leq k-1$, then

$$d = (u_1, \ldots, u_{k-1}) \geq 1.$$

In this case, we define

$$u'_i = \frac{u_i}{d} \qquad \text{for } i = 1, \ldots, k-1,$$

and

$$U' = \max\{|u'_1|, \ldots, |u'_{k-1}|\} \leq \frac{U}{d}.$$

Then $(u'_1, \ldots, u'_{k-1}) = 1$. Consider the linear equation

$$u'_1 v_1 + \cdots + u'_{k-1} v_{k-1} = m'. \tag{12.7}$$

By the induction hypothesis,

$$\begin{aligned} & \text{NSE} \left\{ \begin{array}{l} u_1 v_1 + \cdots + u_{k-1} v_{k-1} = dm' \\ \text{with } |v_1|, \ldots, |v_{k-1}| \leq Q \end{array} \right\} \\ &= \text{NSE } \{(12.7) \quad \text{with } |v_1|, \ldots, |v_{k-1}| \leq Q\} \\ &\leq \frac{(k-2)!(3Q)^{k-2}}{U'}. \end{aligned}$$

If the integer m' can be represented in the form (12.7) with $|v_i| \leq Q$, then

$$|m'| \leq (k-1)U'Q.$$

Since $(d, u_k) = (u_1, \ldots, u_{k-1}, u_k) = 1$ and $\max\{d, |u_k|\} = |u_k| = U$, it follows that

$$\begin{aligned} & \text{NSE} \left\{ \begin{array}{l} u_1 v_1 + \cdots + u_k v_k = m \\ \text{with } |v_1|, \ldots, |v_k| \leq Q \end{array} \right\} \\ &\leq \text{NSE} \left\{ \begin{array}{l} u_1 v_1 + \cdots + u_{k-1} v_{k-1} = dm' \\ \text{with } |v_1|, \ldots, |v_{k-1}| \leq Q \end{array} \right\} \\ &\qquad \times \text{NSE} \left\{ \begin{array}{l} dm' + u_k v_k = m \quad \text{with} \\ |m'|, |v_k| \leq (k-1)U'Q \end{array} \right\} \\ &\leq \frac{(k-2)!(3Q)^{k-2}}{U'} \times \frac{3(k-1)U'Q}{U} \\ &= \frac{(k-1)!(3Q)^{k-1}}{U}. \end{aligned}$$

This completes the proof. □

Theorem 12.2 *Let $k \geq 3$ and let P, Q, and c be real numbers such that*

$$1 \leq P \leq Q \leq cP^{k-1}.$$

Consider the quadratic equation

$$u_1 v_1 + \cdots + u_k v_k = 0 \tag{12.8}$$

in $2k$ variables $u_1, \ldots, u_k, v_1, \ldots, v_k$. Then

$$NSE \left\{ \begin{array}{l} u_1 v_1 + \cdots + u_k v_k = 0 \\ \textit{with } |u_i| \leq P \textit{ and } |v_i| \leq Q \\ \textit{for } i = 1, \ldots k \end{array} \right\} \ll_{k,c} (PQ)^{k-1}.$$

Proof. If $u_1 = \cdots = u_k = 0$, then the number of solutions of (12.8) with $|v_i| \leq Q$ is at most

$$\begin{aligned} (2Q+1)^k &\leq (3Q)^k = 3Q(3Q)^{k-1} \\ &\leq 3cP^{k-1}(3Q)^{k-1} = 3^k c(PQ)^{k-1} \\ &\ll_{k,c} (PQ)^{k-1}. \end{aligned}$$

Suppose that $u_i \neq 0$ for some i. Then

$$1 \leq U = \max\{|u_1|, \ldots, |u_k|\} \leq P.$$

There exists a unique nonnegative integer m such that

$$\frac{P}{2^{m+1}} < U \leq \frac{P}{2^m}. \tag{12.9}$$

The number of equations of the form (12.8) with $|u_i| \leq U \leq P/2^m$ does not exceed

$$\left(\frac{2P}{2^m} + 1\right)^k \leq \left(\frac{3P}{2^m}\right)^k.$$

If

$$(u_1, \ldots, u_k) = 1,$$

then by Lemma 12.3, the number of solutions of each such equation with $|v_i| \leq Q$ is at most

$$\frac{(k-1)!(3Q)^{k-1}}{U} < \frac{(k-1)!2^{m+1}(3Q)^{k-1}}{P}.$$

Therefore, the number of solutions of all equations (12.8) with $(u_1, \ldots, u_k) = 1$ and U in the interval (12.9) is less than

$$\frac{(k-1)!2^{m+1}(3Q)^{k-1}}{P} \left(\frac{3P}{2^m}\right)^k = \frac{6(k-1)!(9PQ)^{k-1}}{2^{(k-1)m}}.$$

Summing over m, we obtain

$$\begin{aligned} \text{NSE}\left\{\begin{array}{c} u_1v_1+\cdots+u_kv_k=0 \\ \text{with } |u_i|\leq P, |v_i|\leq Q, \\ \text{and } (u_1,\ldots,u_k)=1 \end{array}\right\} &< \sum_{m=0}^{\infty}\frac{6(k-1)!(9PQ)^{k-1}}{2^{(k-1)m}} \\ &\leq 8(k-1)!(9PQ)^{k-1}. \end{aligned}$$

If $(u_1,\ldots,u_k)=d$, we define $u_i'=u_i/d$ for $i=1,\ldots,k$. The integers $u_1',\ldots,u_k'$ are relatively prime, and $|u_i'|\leq P/d$. The integers $v_1,\ldots,v_k$ are a solution of equation (12.8) with $|u_i|\leq P$ if and only if $(v_1,\ldots,v_k)$ is a solution of the equation

$$u_1'v_1+\cdots+u_k'v_k=0 \qquad \text{with } |u_i'|\leq P/d.$$

Therefore,

$$\begin{aligned} \text{NSE}\left\{\begin{array}{c} u_1v_1+\cdots+u_kv_k=0 \\ \text{with } |u_i|\leq P, |v_i|\leq Q, \\ \text{and } (u_1,\ldots,u_k)=d \end{array}\right\} &< 8(k-1)!\left(9\left(\frac{P}{d}\right)Q\right)^{k-1} \\ &= \frac{8(k-1)!(9PQ)^{k-1}}{d^{k-1}}. \end{aligned}$$

For $k\geq 3$ we have

$$\sum_{d=1}^{\infty}\frac{1}{d^{k-1}}<1+\int_1^{\infty}\frac{dx}{x^{k-1}}=\frac{k-1}{k-2}\leq 2.$$

Summing over d, we obtain

$$\begin{aligned} \text{NSE}\left\{\begin{array}{c} u_1v_1+\cdots+u_kv_k=0 \\ \text{with } |u_i|\leq P, |v_i|\leq Q, \\ \text{and } u_i\neq 0 \text{ for some } i \end{array}\right\} &< \sum_{d=1}^{\infty}\frac{8(k-1)!(9PQ)^{k-1}}{d^{k-1}} \\ &\leq 16(k-1)!(9PQ)^{k-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\text{NSE}\left\{\begin{array}{c} u_1v_1+\cdots+u_kv_k=0 \\ \text{with } |u_i|\leq P \text{ and } |v_i|\leq Q \end{array}\right\} \\ &< \quad 3^kc(PQ)^{k-1}+16(k-1)!(9PQ)^{k-1} \\ &\ll_{k,c} \quad (PQ)^{k-1}. \end{aligned}$$

This completes the proof. □

Exercises

1. Find all solutions of the linear diophantine equation

$$6v_1 + 10v_2 + 15v_3 = 0 \qquad \text{with } |v_1|, |v_2|, |v_3| \leq 10.$$

   Compare the number of solutions with the upper bound obtained from Lemma 12.3.

2. Find all solutions of the linear diophantine equation

$$6v_1 + 10v_2 + 15v_3 = 1 \qquad \text{with } |v_1|, |v_2|, |v_3| \leq 10.$$

3. Find all solutions of the quadratic equation

$$u_1v_1 + u_2v_2 + u_3v_3 = 0$$

   with $|u_i| \leq 1$ and $|v_i| \leq 1$ for $i = 1, 2, 3$. Compare the number of solutions with the upper bound obtained from Theorem 12.2.

12.3 An Upper Bound for Representations

We can now prove Theorem 12.3, which gives the fundamental upper bound for the number of representations of an integer as the sum of a bounded number of values of polynomials of degree k. We need the following standard result about polynomials.

Lemma 12.4 *Let*

$$f(x) = \sum_{i=0}^{k} a_i x^i$$

be a polynomial of degree k with complex coefficients. Then

$$f(x+u) - f(x) = ug_u(x),$$

where

$$g_u(x) = \sum_{i=0}^{k-1} a_i'(u) x^i$$

is a polynomial of degree $k-1$ with coefficients

$$a_i'(u) = \sum_{j=i+1}^{k} \binom{j}{i} a_j u^{j-i-1}.$$

For any positive number P, if

$$\begin{aligned} |x| &\leq c_1 P, \\ |u| &\leq 2c_1 P, \end{aligned}$$

and

$$|a_i| \leq cP^{k-i} \qquad \text{for } i = 0, 1, \ldots, k,$$

then

$$|a_i'(u)| \leq c(4c_1)^k kP^{k-1-i} \qquad \text{for } i = 0, 1, \ldots, k-1$$

and

$$|g_u(x)| \leq c(2c_1)^{2k} k^2 P^{k-1} \tag{12.10}$$

Proof. This is a purely formal calculation. We have

$$\begin{aligned} f(x+u) - f(x) &= \sum_{j=0}^{k} a_j (x+u)^j - \sum_{j=0}^{k} a_j x^j \\ &= \sum_{j=1}^{k} a_j \sum_{i=0}^{j-1} \binom{j}{i} x^i u^{j-i} \\ &= u \sum_{i=0}^{k-1} \left(\sum_{j=i+1}^{k} \binom{j}{i} a_j u^{j-i-1} \right) x^i \\ &= u g_u(x). \end{aligned}$$

If $|a_i| \leq cP^{k-i}$ and $|u| \leq 2c_1 P$, then

$$\begin{aligned} |a_i'(u)| &\leq \sum_{j=i+1}^{k} \binom{j}{i} |a_j||u|^{j-i-1} \leq \sum_{j=i+1}^{k} 2^j cP^{k-j} (2c_1 P)^{j-i-1} \\ &\leq c(4c_1)^k kP^{k-1-i}. \end{aligned}$$

If also $|x| \leq c_1 P$, then

$$\begin{aligned} |g_u(x)| &\leq \sum_{i=0}^{k-1} |a_i'(u)||x|^i \\ &\leq \sum_{i=0}^{k-1} c(4c_1)^k kP^{k-1-i} (c_1 P)^i \\ &\leq c(2c_1)^{2k} k^2 P^{k-1}. \end{aligned}$$

This completes the proof. □

Theorem 12.3 *Let $\{s(k)\}_{k=1}^{\infty}$ be the sequence of integers defined recursively by $s(1) = 1$ and*

$$s(k) = 8k2^{[\log_2 s(k-1)]} \qquad \text{for } k \geq 2. \tag{12.11}$$

Let $c \geq 1$. For $j = 1, \ldots, s(k)$, let

$$f_j(x) = \sum_{i=0}^{k} a_{ij} x^i$$

be a sequence of polynomials with complex coefficients such that

$$|a_{kj}| \leq c \qquad \text{for } j = 1, \ldots, s(k).$$

Choose $P \geq 1$ such that

$$|a_{ij}| \leq cP^{k-i} \qquad \text{for } i = 0, 1, \ldots, k-1 \text{ and } j = 1, \ldots, s(k). \tag{12.12}$$

Let $c_1 \geq 1$. For every complex number z,

$$NSE \left\{ \begin{array}{l} \sum_{j=1}^{s(k)} f_j(x_j) = z \quad \text{with } x_j \in \mathbf{Z} \\ \text{and } |x_j| \leq c_1 P \text{ for } j = 1, \ldots, s(k) \end{array} \right\} \ll_{k,c,c_1} P^{s(k)-k}. \tag{12.13}$$

Proof. The proof is by induction on the degree k of the polynomials.

For $k = 1$ we have $s(1) = 1$ and $f_1(x) = a_{11}x + a_{01}$. For any number z, there exists at most one integer x_1 such that $f_1(x_1) = z$, and so

$$\text{NSE} \left\{ \begin{array}{l} f_1(x_1) = z \quad \text{with } x_1 \in \mathbf{Z} \\ \text{and } |x_1| \leq c_1 P \end{array} \right\} \leq 1 = P^{s(1)-1}.$$

Let $k \geq 2$, and assume that the theorem holds for $s' = s(k-1)$ polynomials of degree $k-1$. Define

$$t = t(k) = [\log_2 s'] + 2$$

and

$$s = s(k) = 2k2^t = 8k2^{[\log_2 s(k-1)]}.$$

Since $[x] \leq x < [x] + 1$ for every real number x, we have

$$s' = 2^{\log_2 s'} < 2^{[\log_2 s'] + 1} = 2^{t-1}.$$

Consider the weighted set $(X, w_X))$, where

$$X = \{x \in \mathbf{Z} : |x| \leq c_1 P\}$$

and $w_X(x) = 1$ for all $x \in X$. For $j = 1, \ldots, s$ we have the weighted sets

$$F_j = \{f_j(x) : x \in X\} = \{f_j(x) : |x| \leq c_1 P\}$$

with weights

$$w_{F_j}^{(f_j)}(z) = \text{NSE}\ \{f_j(x) = z : |x| \leq c_1 P\}.$$

Let S be the weighted sumset

$$S = F_1 + \cdots + F_s.$$

Then

$$w_S(z) = \mathrm{NSE}\left\{\sum_{j=1}^{s} f_j(x_j) = z \quad \text{with } |x_j| \leq c_1 P\right\}.$$

For

$$m = \frac{s}{2} = k2^t,$$

we consider the weighted sumsets

$$B_1 = F_1 + \cdots + F_m$$

and

$$B_2 = F_{m+1} + \cdots + F_{2m},$$

and the weighted difference sets

$$D_1 = B_1 - B_1 = \left\{\sum_{j=1}^{m}(f_j(y_j) - f_j(x_j)) : |x_j|, |y_j| \leq c_1 P\right\}$$

and

$$D_2 = B_2 - B_2 = \left\{\sum_{j=m+1}^{2m}(f_j(y_j) - f_j(x_j)) : |x_j|, |y_j| \leq c_1 P\right\}.$$

Applying Lemma 12.1 to $S = B_1 + B_2$, we obtain

$$w_S(z) \leq \frac{1}{2}\left(w_{D_1}(0) + w_{D_2}(0)\right).$$

For $j = 1, \ldots, s$, let

$$f_j(x + u) - f_j(x) = u g_{j,u}(x),$$

where $g_{j,u}(x)$ is the polynomial of degree $k-1$ constructed in Lemma 12.4. We can use our result on quadratic equations and weighted sets (Theorem 12.2) to obtain upper bounds for the weights $w_{D_1}(0)$ and $w_{D_2}(0)$. If $|x_j|, |y_j| \leq c_1 P$ and $u_j = y_j - x_j$, then $|u_j| \leq |x_j| + |y_j| \leq 2c_1 P$. It follows that

$$\begin{aligned}
w_{D_1}(0) &= \mathrm{NSE}\left\{\begin{array}{l}\sum_{j=1}^{m}(f_j(y_j) - f_j(x_j)) = 0 \\ \text{with } |x_j|, |y_j| \leq c_1 P\end{array}\right\} \\
&\leq \mathrm{NSE}\left\{\begin{array}{l}\sum_{j=1}^{m}(f_j(x_j + u_j) - f_j(x_j)) = 0 \\ \text{with } |x_j| \leq c_1 P \text{ and } |u_j| \leq 2c_1 P\end{array}\right\} \\
&= \mathrm{NSE}\left\{\begin{array}{l}\sum_{j=1}^{m} u_j g_{j,u_j}(x_j) = 0 \\ \text{with } |x_j| \leq c_1 P \text{ and } |u_j| \leq 2c_1 P\end{array}\right\} \\
&= \sum_{|u_1|,\ldots,|u_m| \leq 2c_1 P} \mathrm{NSE}\left\{\begin{array}{l}\sum_{j=1}^{m} u_j g_{j,u_j}(x_j) = 0 \qquad \text{with} \\ |x_j| \leq c_1 P \text{ for } j = 1, \ldots, m\end{array}\right\}.
\end{aligned}$$

Similarly,

$$w_{D_2}(0) \leq \sum_{|u_{m+1}|,\ldots,|u_{2m}|\leq 2c_1P} \mathrm{NSE}\left\{\begin{array}{ll} \sum_{j=m+1}^{2m} u_j g_{j,u_j}(x_j) = 0 & \text{with} \\ |x_j| \leq c_1 P \text{ for } j = m+1,\ldots,2m \end{array}\right\}.$$

For $j = 1, \ldots, m$, we fix integers u_j with $|u_j| \leq 2c_1P$, and consider the weighted sets

$$G_j = \{g_{j,u_j}(x) : |x| \leq c_1 P\}$$

and

$$G'_j = u_j * \{g_{j,u_j}(x) : |x| \leq c_1P\} = \{u_j g_{j,u_j}(x) : |x| \leq c_1 P\},$$

with weights

$$w_{G_j}(z) = w_{G'_j}(u_j z) = \mathrm{NSE}\ \{g_{j,u_j}(x) = z : |x| \leq c_1P\}.$$

Recall that $m = k2^t$. For $q = 1, \ldots, 2^t$, we define the weighted sets

$$\begin{aligned} B'_q &= G'_{(q-1)k+1} + G'_{(q-1)k+2} + \cdots + G'_{qk}, \\ D'_q &= 2^{t-1}B'_q - 2^{t-1}B'_q, \end{aligned}$$

and

$$S'_1 = \sum_{j=1}^{m} G'_j = \sum_{q=1}^{2^t} B'_q.$$

Then

$$w_{S'_1}(0) = \mathrm{NSE}\left\{\sum_{j=1}^{m} u_j g_{j,u_j}(x_j) = 0 \quad \text{with } |x_j| \leq c_1 P\right\}.$$

By Lemma 12.2,

$$w_{S'_1}(0) \leq \frac{1}{2^t}\sum_{q=1}^{2^t} w_{D'_q}(0).$$

We can express the difference set D'_q as follows:

$$\begin{aligned} D'_q &= 2^{t-1}B'_q - 2^{t-1}B'_q \\ &= 2^{t-1}\sum_{r=1}^{k} G'_{(q-1)k+r} - 2^{t-1}\sum_{r=1}^{k} G'_{(q-1)k+r} \\ &= \sum_{r=1}^{k} u_{(q-1)k+r} * \left(2^{t-1}G_{(q-1)k+r} - 2^{t-1}G_{(q-1)k+r}\right) \\ &= \sum_{r=1}^{k} u_{(q-1)k+r} * V_{(q-1)k+r}, \end{aligned}$$

where
$$V_{(q-1)k+r} = 2^{t-1}G_{(q-1)k+r} - 2^{t-1}G_{(q-1)k+r}.$$

Let $v \in V_{(q-1)k+r}$. By Lemma 12.4, if $|x| \leq c_1 P$, then

$$|g_{(q-1)k+r,u_{(q-1)k+r}}(x)| \leq c(2c_1)^{2k}k^2 P^{k-1},$$

and so
$$|v| \leq c(2c_1)^{2k}k^2 2^t P^{k-1}. \tag{12.14}$$

We shall use the induction hypothesis for polynomials of degree $k-1$ to obtain an upper bound for the weight of v. Let

$$g_u(x) = g_{(q-1)k+r,u_{(q-1)k+r}}(x) = \sum_{i=0}^{k-1} a_i'(u)x^i.$$

By Lemma 12.4, we have

$$|a_i'(u)| \leq c(4c_1)^k k P^{k-1-i}$$

for $i = 0, 1, \ldots, k-1$. Since $s' = s(k-1)$, for every number z' we have

$$\text{NSE}\left\{\begin{array}{l} \sum_{j=1}^{s'} g(x_j) = z' \\ \text{with } |x_j| \leq c_1 P \text{ for } j = 1, \ldots, s' \end{array}\right\} \ll_{k,c,c_1} P^{s'-k+1}.$$

Since $s' < 2^{t-1}$, we obtain the following upper bound for the weight of v:

$$\begin{aligned}
w_{V_{(q-1)k+r}}(v) &= \text{NSE}\left\{\begin{array}{l} v = \sum_{q=1}^{2^{t-1}} g_u(x_q) - \sum_{q=1}^{2^{t-1}} g_u(x_q') \\ \text{with } |x_q|, |x_q'| \leq c_1 P \text{ for } q = 1, \ldots, 2^{t-1} \end{array}\right\} \\
&= \text{NSE}\left\{\begin{array}{l} \sum_{q=1}^{s'} g_u(x_q) = \\ v + \sum_{q=1}^{2^{t-1}} g_u(x_q') - \sum_{q=s'+1}^{2^{t-1}} g_u(x_q) \\ \text{with } |x_q|, |x_q'| \leq c_1 P \text{ for } q = 1, \ldots, 2^{t-1} \end{array}\right\} \\
&= \sum_{\substack{|x_1'|,\ldots,|x_{2^{t-1}}'|, \\ |x_{s'+1}|,\ldots,|x_{2^{t-1}}| \leq c_1 P}} \text{NSE}\left\{\begin{array}{l} \sum_{q=1}^{s'} g_u(x_q) = v + \\ \sum_{q=1}^{2^{t-1}} g_u(x_q') - \sum_{q=s'+1}^{2^{t-1}} g_u(x_q) \\ \text{with } |x_q| \leq c_1 P \text{ for } q = 1, \ldots, s' \end{array}\right\} \\
&\ll_{k,c,c_1} \sum_{\substack{|x_1'|,\ldots,|x_{2^{t-1}}'|, \\ |x_{s'+1}|,\ldots,|x_{2^{t-1}}| \leq c_1 P}} P^{s'-(k-1)} \\
&\ll_{k,c,c_1} P^{2^t - s'} P^{s'-k+1} \\
&\ll_{k,c,c_1} P^{2^t-k+1}.
\end{aligned}$$

Therefore, there exists a constant $c' = c(k, c, c_1)$ such that $w_{V_j}(v) \leq c' P^{2^t-k+1}$ for all $j = 1, \ldots, m$ and $v \in V_j$.

Let U be the weighted set of all integers u such that $|u| \leq 2c_1 P$ and $w_U(u) = 1$ for all $u \in U$. Let V be the weighted set of all integers v that satisfy inequality (12.14) and have constant weight $w_V(v) = cP^{2^t-k+1}$. We can now find an upper bound for the weights $w_{D_1}(0)$ and $w_{D_2}(0)$:

$$
\begin{aligned}
w_{D_1}(0) &\leq \sum_{|u_1|,\ldots,|u_m| \leq 2c_1 P} w_{S_1'(0)} \\
&\leq \sum_{|u_1|,\ldots,|u_m| \leq 2c_1 P} \frac{1}{2^t} \sum_{q=1}^{2^t} w_{D_q'}(0) \\
&\leq \sum_{|u_1|,\ldots,|u_m| \leq 2c_1 P} \frac{1}{2^t} \\
&\quad \times \sum_{q=1}^{2^t} \left(c' P^{2^t-k+1} \right)^k \mathrm{NSE} \left\{ \begin{array}{l} \sum_{r=1}^k u_{(q-1)k+r} v_{(q-1)k+r} = 0 \\ \text{with } v_{(q-1)k+r} \in V_{(q-1)k+r} \end{array} \right\} \\
&\ll_{k,c.c_1} \frac{P^{m-k^2+k}}{2^t} \\
&\quad \times \sum_{q=1}^{2^t} \sum_{u_1,\ldots,u_m \in U} \mathrm{NSE} \left\{ \begin{array}{l} \sum_{r=1}^k u_{(q-1)k+r} v_{(q-1)k+r} = 0 \\ \text{with } v_{(q-1)k+r} \in V \end{array} \right\} \\
&= \frac{P^{m-k^2+k}}{2^t} \\
&\quad \times \sum_{q=1}^{2^t} \sum_{\substack{u_1,\ldots,u_{(q-1)k}, \\ u_{qk+1},\ldots,u_m \in U}} \mathrm{NSE} \left\{ \begin{array}{l} \sum_{r=1}^k u_{(q-1)k+r} v_{(q-1)k+r} = 0 \\ \text{with } v_{(q-1)k+r} \in V \text{ and} \\ u_{(q-1)k+1}, \ldots, u_{qk} \in U \end{array} \right\} \\
&\ll_{k,c,c_1} \frac{P^{m-k^2+k} P^{m-k}}{2^t} \\
&\quad \times \sum_{q=1}^{2^t} \mathrm{NSE} \left\{ \begin{array}{l} \sum_{r=1}^k u_{(q-1)k+r} v_{(q-1)k+r} = 0 \\ \text{with } v_{(q-1)k+r} \in V_{(q-1)k+r} \text{ and} \\ u_{(q-1)k+1}, \ldots, u_{qk} \in U \end{array} \right\} \\
&\ll_{k,c,c_1} P^{s-k^2} \mathrm{NSE} \left\{ \begin{array}{l} \sum_{r=1}^k u_r v_r = 0 \\ \text{with } v_r \in V \text{ and } u_r \in U \end{array} \right\} \\
&\ll_{k,cc_1} P^{s-k^2} (P^k)^{k-1} \qquad \text{(by Theorem 12.2)} \\
&\ll_{k,c,c_1} P^{s-k}.
\end{aligned}
$$

Similarly,

$$w_{D_2}(0) \ll_{k,c,c_1} P^{s-k}.$$

Therefore,

$$w_S(n) \leq \frac{1}{2}(w_{D_1}(0) + w_{D_2}(0)) \ll_{k,c,c_1} P^{s-k}.$$

This completes the proof. □

Exercises

1. Compute $s(k)$ for $k = 1, 2, 3, 4, 5$.
2. Prove that
$$4^{k-1}k! < s(k) \leq 8^{k-1}k!$$
for $k \geq 2$.

12.4 Waring's Problem for Sequences of Polynomials

In Chapter 11 we applied a special case of Theorem 12.3 to prove Waring's problem for a polynomial. In this section we show how the full strength of Theorem 12.3 yields a generalization of Waring's problem to finite sequences of polynomials. Let $c \geq 1$. For $j = 1, \ldots, s$, let $f_j(x)$ be an integer-valued polynomial of degree k whose leading coefficient a_{kj} satisfies the inequality $0 < a_{kj} \leq c$. We consider the sequence

$$\mathcal{F} = \{f_j(x)\}_{j=1}^{s}.$$

We shall prove that there exist integers $s(k)$ and $h(k)$ and a positive number $\delta(k, c)$ such that if $s \geq s(k)$, then the set

$$S = \{f_1(x_1) + \cdots + f_s(x_s) : x_1, \ldots, x_s \in \mathbf{N}_0\}$$

has lower asymptotic density $d_L(S) \geq \delta(k, c) > 0$, and if $s \geq h(k)$, then S is eventually coincides with a union of congruence classes.

We define the representation functions $r_{\mathcal{F}}(n)$ and $R_{\mathcal{F}}(N)$ by

$$r_{\mathcal{F}}(n) = \text{NSE} \left\{ \begin{array}{l} f_1(x_1) + \cdots + f_s(x_s) = n \\ \text{with } x_1, \ldots, x_s \in \mathbf{N}_0 \end{array} \right\}$$

and

$$R_{\mathcal{F}}(N) = \sum_{0 \leq n \leq N} r_{\mathcal{F}}(n).$$

Lemma 12.5 *Let $c \geq 1$. Let $\mathcal{F} = \{f_j(x)\}_{j=1}^{s}$ be a sequence of integer-valued polynomials of degree k, and let a_{kj} be the leading coefficient of $f_j(x)$. We assume that*

$$0 < a_{kj} \leq c$$

for $j = 1, \ldots, s$. If N is sufficiently large, then

$$R_{\mathcal{F}}(N) > \frac{1}{2}\left(\frac{2N}{3cs}\right)^{s/k}. \tag{12.15}$$

Proof. Define $x^*(f_j)$ by (11.4) for $j = 1, \ldots, s$. If the integers x_j satisfy the inequalities

$$x^*(f_j) \leq x_j \leq \left(\frac{2N}{3cs}\right)^{1/k},$$

then, by Lemma 11.4,

$$0 \leq f_j(x_j) \leq \frac{3a_{kj}x_j^k}{2} \leq \frac{3c}{2}\left(\frac{2N}{3cs}\right) = \frac{N}{s}$$

and

$$0 \leq f_1(x_1) + \cdots + f_s(x_s) \leq N.$$

Therefore,

$$R_{\mathcal{F}}(N) > \left(\left(\frac{2N}{3cs}\right)^{1/k} - x^*(f) - 1\right)^s \geq \frac{1}{2}\left(\frac{2N}{3cs}\right)^{s/k}$$

for N sufficiently large. This proves (12.15). □

Lemma 12.6 *Let $\mathcal{F} = \{f_j(x)\}_{j=1}^s$ be a sequence of integer-valued polynomials of degree k, and let a_{kj} be the leading coefficient of $f_j(x)$. Let $c \geq 1$. We assume that*

$$0 < a_{kj} \leq c$$

and that $A(f_j) = \{f_j(x) : x \in \mathbf{N}_0\}$ is a strictly increasing sequence of nonnegative integers for $j = 1, \ldots, s$. There exists a number $N_1(\mathcal{F})$ such that if $N \geq N_1(\mathcal{F})$ and $x_1, \ldots, x_s$ are nonnegative integers with

$$\sum_{j=1}^{s} f(x_j) \leq N,$$

then

$$x_j \leq (4k!N)^{1/k} \qquad \text{for } j = 1, \ldots, s.$$

Proof. The proof is the same as the proof of Lemma 11.5. Recall that $k!a_{kj} \geq 1$ by Exercise 6 in Section 11.1. Define $x^*(f_j)$ by (11.4) for $j = 1, \ldots, s$, and $x^*(\mathcal{F}) = \max\{x^*(f_1), \ldots, x^*(f_s)\}$. Let

$$N_1(\mathcal{F}) = \frac{x^*(\mathcal{F})^k}{2k!}. \tag{12.16}$$

If $N \geq N_1(\mathcal{F})$, $1 \leq \ell \leq s$, and $x_\ell > (2k!N)^{1/k} \geq x^*(\mathcal{F}) \geq x^*(f_\ell)$, then

$$f_\ell(x_\ell) \geq \frac{a_{k\ell}x_\ell^k}{2} > k!a_k N \geq N,$$

and so

$$\sum_{j=1}^{s} f(x_j) \geq f(x_\ell) \geq f(x_\ell) > N.$$

It follows that if $x_1, \ldots, x_s$ are nonnegative integers such that

$$\sum_{j=1}^{s} f(x_j) \leq N,$$

then

$$x_j \leq (2k!N)^{1/k} \qquad \text{for } j = 1, \ldots, s.$$

This completes the proof. □

Theorem 12.4 *For any positive integer k and real number $c \geq 1$, there exists a number $\delta(k,c) > 0$ with the following property: If $s = s(k)$ is the integer defined by (12.11), and if $\mathcal{F} = \{f_j(x)\}_{j=1}^{s}$ is a sequence of integer-valued polynomials of degree k whose leading coefficients a_{kj} satisfy*

$$0 < a_{kj} \leq c,$$

then the sumset

$$B = \{f_1(x_1) + \cdots + f_s(x_s) : x_1, \ldots, x_s \in \mathbf{N}_0\}$$

has lower asymptotic density

$$d_L(B) \geq \delta(k,c) > 0.$$

Proof. Replacing the polynomial $f_j(x)$ with $f_j(x + x_0)$ for a sufficiently large integer x_0, we can assume that $\{f_j(x) : x \in \mathbf{N}_0\}$ is a strictly increasing sequence of nonnegative integers for $j = 1, \ldots, s$.

Define $N_1(\mathcal{F})$ by (12.16). Choose $N_2(\mathcal{F})$ sufficiently large that for $N \geq N_2(\mathcal{F})$ and $P = N^{1/k}$, we have

$$|a_{ij}| \leq cP^{k-i} \qquad \text{for } i = 0, 1, \ldots, k-1,$$

and so Theorem 12.3 applies to the polynomials in the sequence $\mathcal{F}$.

Let $N(\mathcal{F}) = \max\{N_1(\mathcal{F}), N_2(\mathcal{F}))$ and $c_1 = (2k!)^{1/k}$. By Lemma 12.6, if $N \geq N(\mathcal{F})$ and $x_1, \ldots, x_s$ are nonnegative integers such that

$$f_1(x_1) + \cdots + f_s(x_s) \leq N,$$

then $x_j \leq c_1 P$ for $j = 1, \ldots, s$. Therefore, if $0 \leq n \leq N$, then

$$\begin{aligned} r_{\mathcal{F}}(n) &= \text{NSE}\left\{ \begin{array}{l} f_1(x_1) + \cdots + f_s(x_s) = n \\ \text{with } x_j \in \mathbf{N}_0 \text{ for } j = 1, \ldots, s(k) \end{array} \right\} \\ &= \text{NSE}\left\{ \begin{array}{l} f_1(x_1) + \cdots + f_s(x_s) = n \\ \text{with } 0 \leq x_j \leq c_1 P \text{ for } j = 1, \ldots, s(k) \end{array} \right\} \\ &\ll_{k,c} P^{s-k} \end{aligned}$$

by Theorem 12.3. Let $B(n)$ be the counting function of the set B. We have

$$\begin{aligned} R_{\mathcal{F}}(N) &= \sum_{n=0}^{N} r_{\mathcal{F}}(n) = \sum_{\substack{n=0 \\ n \in B}}^{N} r_{\mathcal{F}}(n) \\ &\ll_{k,c} B(N)P^{s-k} = \frac{B(N)P^s}{N}. \end{aligned}$$

By Lemma 12.5,

$$R_{\mathcal{F}}(N) > \frac{1}{2}\left(\frac{2}{3cs}\right)^{s/k} P^s.$$

It follows that $B(N)/N \gg_{k,c} 1$. This completes the proof. □

We say that sets of integers A and B *eventually coincide* if there exists a number n_0 such that $n \in A$ if and only if $n \in B$ for all $n \geq n_0$. By Theorem 12.4, the set of sums of $s(k)$ integer-valued polynomials of degree k has positive lower asymptotic density, but not necessarily a rich arithmetic structure. For example, sets of positive density can have arbitrarily large gaps between consecutive elements. We shall prove that there exists a number $h = h(k, c)$ such that the set of sums of $h(k, c)$ integer-valued polynomials of degree k with positive leading coefficients not exceeding c has bounded gaps, and, moreover, eventually coincides with a union of congruence classes. The proof of this result requires a *deus ex machina* in the form of a theorem of Kneser on the asymptotic density of sumsets. We do not prove Kneser's theorem in this book, but this application of Kneser's theorem gives a generalization of Waring's problem that is too beautiful to resist.

For $i = 1, \ldots, d$, let B_i be a set of integers with lower asymptotic density $d_L(B_i) = \beta_i$, and let $S = B_1 + \cdots + B_d$. Kneser's theorem states that if $d_L(S) < \beta_1 + \cdots + \beta_d$, then there is a modulus $m \geq 1$ such that the sumset S eventually coincides with a union of congruence classes modulo m.

Theorem 12.5 *Let k be a positive integer and $c \geq 1$. There exists a positive integer $h = h(k, c)$ with the following property: Let $\mathcal{F} = \{f_j(x)\}_{j=1}^{h}$ be a sequence of integer-valued polynomials of degree k such that the leading*

coefficient a_{kj} *of* $f_j(x)$ *satisfies the inequality* $0 < a_{kj} \leq c$ *for* $j = 1, \ldots, h$. *There exists a positive integer* m *such that the sumset*

$$S = \{f_1(x_1) + \cdots + f_h(x_h) : x_j \in \mathbf{N}_0 \quad \text{for } j = 1, \ldots, h\}$$

eventually coincides with a union of congruence classes modulo m.

Proof. Let $s = s(k)$ be the positive integer constructed in Theorem 12.3 and let $\delta = \delta(k, c)$ be the positive number constructed in Theorem 12.4. We define

$$d = \left[\frac{1}{\delta}\right] + 1$$

and

$$h = h(k, c) = ds.$$

Let $\mathcal{F} = \{f_j(x)\}_{j=1}^{h}$ be a sequence of integer-valued polynomials of degree k whose leading coefficients are positive and not greater than c. For $i = 1, \ldots, d$, let $\mathcal{F}_i = \{f_{(i-1)s+j}(x)\}_{j=1}^{s}$. By Theorem 12.4, the sumset

$$B_i = \left\{\sum_{j=1}^{s} f_{(i-1)s+j}(x_j) : x_j \in \mathbf{N}_0\right\}$$

has lower asymptotic density $d_L(B_i) \geq \delta > 0$. Since

$$S = B_1 + \cdots + B_d = \left\{\sum_{j=1}^{h} f_j(x_j) : x_j \in \mathbf{N}_0\right\}$$

and

$$\sum_{i=1}^{d} d_L(B_i) \geq \delta d = \delta\left(\left[\frac{1}{\delta}\right] + 1\right) > 1 \geq d_L(S),$$

Kneser's theorem implies that S eventually coincides with a union of congruence classes modulo m for some positive integer m. □

12.5 Notes

> This proof, so exquisitely elementary, will undoubtedly seem very complicated to you. But it will take you only two to three weeks' work with pencil and paper to understand and digest it completely. It is by conquering difficulties of just this sort, that the mathematician grows and develops.

A. Ya. Khinchin [78]

The proof to which Khinchin refers is Linnik's elementary proof of Waring's problem. It is the "third pearl" in Khinchin's famous book *Three Pearls of Number Theory* [78]. The quotation is the last paragraph in the book.

Theorem 12.3 generalizes a result of Linnik for sums of one polynomial to sums of a sequence of polynomials. Linnik's result provides the essential upper bound in his solution of Waring's problem.

Often, theorems in number theory and, in particular, variants of Waring's problem, are first proved analytically, and only later are elementary proofs discovered. Theorem 12.4, due to Nathanson, is an unusual example of a result that was first proved by elementary methods.

For a proof of Kneser's theorem [79] on the asymptotic density of sumsets, see Halberstam and Roth [48] and Nathanson [108].

13
Liouville's Identity

13.1 A Miraculous Formula

In a series of eighteen papers published between 1858 and 1865, Liouville introduced a strange and powerful method into elementary number theory. In this chapter we prove an important identity of Liouville. We shall apply it in Chapter 14 to obtain theorems about the number of representations of an integer as a sum of an even number of squares. This is our second problem in additive number theory.

Recall that a function $f(x)$ is called *even* if $f(-x) = f(x)$ for all x. A function $f(x)$ is called *odd* if $f(-x) = -f(x)$ for all x. If $f(x)$ is odd, then $f(0) = -f(0)$, and so $f(0) = 0$.

The function $F(x, y, z)$ is odd in the variable x if $F(-x, y, z) = -F(x, y, z)$, and even in the pair of variables (y, z) if $F(x, -y, -z) = F(x, y, z)$. If $F(x, y, z)$ is odd in the variable y and also odd in the variable z, then $F(x, y, z)$ is even in the pair of variables (y, z). For example, the function $F(x, y, z) = xyz$ is odd in the variable x and even in the pair of variables (y, z).

In this and the following chapter, u, v, and w denote integers, and d, δ, and ℓ denote positive integers. The notation

$$\sum_{u^2+d\delta=n}$$

means the sum over all ordered triples (u, d, δ) such that $u^2 + d\delta = n$. For example,

$$\begin{aligned}\sum_{u^2+d\delta=3} G(u,d,\delta) &= G(0,1,3) + G(0,3,1) + G(1,1,2) \\ &\quad + G(1,2,1) + G(-1,1,2) + G(-1,2,1).\end{aligned}$$

We define the symbol $\{T(\ell)\}_{n=\ell^2}$ as follows:

$$\{T(\ell)\}_{n=\ell^2} = \begin{cases} 0 & \text{if } n \text{ is not a square,} \\ T(\ell) & \text{if } n \text{ is a square and } n = \ell^2. \end{cases}$$

Liouville's fundamental identity is the following.

Theorem 13.1 (Liouville) *Let $F(x, y, z)$ be a function defined on the set of all triples (x, y, z) of integers such that $F(x, y, z)$ is odd in the variable x and even in the pair of variables (y, z). For every positive integer n,*

$$\begin{aligned} 2 &\sum_{u^2+d\delta=n} F(\delta - 2u, u + d, 2u + 2d - \delta) \\ &= \sum_{u^2+d\delta=n} F(d + \delta, u, d - \delta) + \{2T_1(\ell) - T_2(\ell)\}_{n=\ell^2}, \end{aligned}$$

where

$$T_1(\ell) = \sum_{j=1}^{2\ell-1} F(j, \ell, j)$$

and

$$T_2(\ell) = \sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j).$$

For example, there are six triples (u, d, δ) such that $u^2+d\delta = 3$. Liouville's formula for $n = 3$ asserts that

$$\begin{aligned} &2(F(3,1,-1) + F(1,3,5) + F(0,2,2) + F(-1,3,5) + F(4,0,-2) + F(3,1,1)) \\ &= F(4,0,2) + F(4,0,-2) + F(3,1,1) + F(3,1,-1) + F(3,-1,1) \\ &\quad + F(3,-1,-1). \end{aligned}$$

It is easy to check this identity using only the parity properties of the function $F(x, y, z)$.

We shall prove Theorem 13.1 in Section 13.4.

Liouville's identity is very general, and we can specialize it in many ways. Here is an example.

Theorem 13.2 *Let $f(y)$ be an odd function. For every positive integer n,*

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u f(u + d) = \left\{(-1)^{\ell-1} \ell f(\ell)\right\}_{n=\ell^2}.$$

Proof. We define the function

$$F(x,y,z) = \begin{cases} 0 & \text{if } x \text{ or } z \text{ is even,} \\ (-1)^{(x+z)/2} f(y) & \text{if } x \text{ and } z \text{ are odd.} \end{cases}$$

Then $F(x,y,z)$ is an odd function of each of the variables x, y, and z, hence an even function of the pair of variables (y,z). If x,y,z are integers and δ is even, then $\delta - 2x$ is even, and so $F(\delta - 2x, y, z) = 0$.

We shall apply Theorem 13.1 to the function $F(x,y,z)$. The left side of Liouville's identity is

$$\begin{aligned} & 2 \sum_{u^2+d\delta=n} F(\delta - 2u, u+d, 2u+2d-\delta) \\ &= 2 \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} F(\delta - 2u, u+d, 2u+2d-\delta) \\ &= 2 \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^d f(u+d) \\ &= 2 \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^{d\delta} f(u+d) \\ &= 2 \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^{n-u^2} f(u+d) \\ &= 2(-1)^n \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^u f(u+d). \end{aligned}$$

The right side of Liouville's identity is

$$\sum_{u^2+d\delta=n} F(d+\delta, u, d-\delta) + \{2T_1(\ell) - T_2(\ell)\}_{n=\ell^2}.$$

If $u^2 + d\delta = n$, then also $(-u)^2 + d\delta = n$, and the map

$$(u,d,\delta) \mapsto (-u,d,\delta) \tag{13.1}$$

is an involution[1] on the set of solutions of the equation $u^2 + d\delta = n$. Then

$$\begin{aligned} \sum_{u^2+d\delta=n} F(d+\delta, u, d-\delta) &= \sum_{u^2+d\delta=n} F(d+\delta, -u, d-\delta) \\ &= -\sum_{u^2+d\delta=n} F(d+\delta, u, d-\delta), \end{aligned}$$

[1] An *involution* on a set X is a map $\alpha : X \to X$ such that α^2 is the identity map.

since $F(x,y,z)$ is an odd function of y. Therefore,

$$\sum_{u^2+d\delta=n} F(d+\delta, u, d-\delta) = 0.$$

If $n = \ell^2$, then

$$\begin{aligned} T_1(\ell) &= \sum_{j=1}^{2\ell-1} F(j,\ell,j) = \sum_{\substack{1\leq j\leq 2\ell-1 \\ j\equiv 1 \pmod 2}} F(j,\ell,j) \\ &= \sum_{i=1}^{\ell} F(2i-1,\ell,2i-1) = -\sum_{j=1}^{\ell} f(\ell) \\ &= -\ell f(\ell) \end{aligned}$$

and

$$T_2(\ell) = \sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j) = 0.$$

Therefore,

$$\begin{aligned} 2\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u f(u+d) &= (-1)^n\{-\ell f(\ell)\}_{n=\ell^2} \\ &= \{(-1)^{\ell-1}\ell f(\ell)\}_{n=\ell^2}. \end{aligned}$$

This completes the proof.

Exercises

1. Let $F(x,y,z)$ be a function that is odd in x and even in (y,z). Write out Liouville's formula in the case $n=4$, and confirm it directly using only the parity properties of $F(x,y,z)$.

2. Prove that for every positive integer n the diophantine equation

$$u^2 + vw = n$$

has infinitely many solutions in integers u, v, w, but only finitely many solutions in integers with $v \geq 1$ and $w \geq 1$.

13.2 Prime Numbers and Quadratic Forms

A *quadratic form* is a homogeneous polynomial of degree two. The quadratic form $Q(x,y,\ldots,z)$ *represents* the integer n if there exist integers $a,b,\ldots,c$

such that $Q(a, b, \ldots, c) = n$. A *binary quadratic form* is a quadratic form in two variables. A *ternary quadratic form* is a quadratic form in three variables. In this section we apply Theorem 13.2 to obtain classical theorems about the representation of prime numbers by the binary quadratic forms $x^2 + y^2$ and $x^2 + 2y^2$.

We begin with some results about divisors. Recall that a positive integer d is called a *divisor* of the positive integer n if there exists an integer δ such that $n = d\delta$. The integer δ is called the *conjugate divisor* of d. The *divisor function* $\sigma(n)$ is the sum of the divisors of n, that is, the arithmetic function defined by

$$\sigma(n) = \sum_{d|n} d.$$

We denote by $\sigma^*(n)$ the sum of the divisors of n whose conjugate divisors are odd. For example, $\sigma(10) = 1+2+5+10 = 17$ and $\sigma^*(10) = 2+10 = 12$. If p is an odd prime, then $\sigma(p) = \sigma^*(p) = p + 1$.

Lemma 13.1 *Let n be an odd positive integer. Then $\sigma(n)$ is odd if and only if n is a square.*

Proof. Let

$$n = \prod_{p|n} p^{v_p}$$

be the unique factorization of n as a product of odd prime numbers. The positive integer d divides n if and only if d can be written in the form

$$d = \prod_{p|n} p^{u_p},$$

where

$$0 \leq u_p \leq v_p,$$

and so

$$\begin{aligned}
\sigma(n) &= \prod_{p|n} \sum_{u_p=0}^{v_p} p^{u_p} \\
&\equiv \prod_{p|n} (u_p + 1) \pmod{2} \\
&\equiv 1 \pmod{2}
\end{aligned}$$

if and only if u_p is even for all p, that is, $u_p = 2w_p$ and

$$n = \prod_{p|n} p^{v_p} = \left(\prod_{p|n} p^{w_p} \right)^2$$

is a square. This completes the proof.

Lemma 13.2 *If $n = 2^k m$, where $k \geq 0$ and m is odd, then $\sigma^*(n) = 2^k\sigma(m)$. If $\sigma^*(n)$ is odd, then n is the square of an odd integer.*

Proof. Let d be a divisor of n. If the conjugate divisor $\delta = n/d$ is odd, then 2^k must divide d, and so $d = 2^k d'$ for some integer d'. Then

$$2^k m = n = d\delta = 2^k d'\delta,$$

and d' is a divisor of m. Conversely, if d' is any divisor of m, then $2^k d'$ is a divisor of n whose conjugate divisor m/d' is odd. Therefore,

$$\sigma^*(n) = 2^k \sum_{d'|m} d' = 2^k\sigma(m).$$

If $\sigma^*(n)$ is odd, then $k = 0$ and $n = m$ is odd. It follows that $\sigma^*(n) = \sigma(m) = \sigma(n)$ is odd, and so n is a square by Lemma 13.1. This completes the proof.

Lemma 13.3 *For every positive integer n,*

$$\sigma^*(n) = 2\sum_{1\leq u<\sqrt{n}} (-1)^{u-1}\sigma^*(n-u^2) + \{(-1)^{n-1}n\}_{n=\ell^2}.$$

Proof. We apply Theorem 13.2 to the odd function $f(y) = y$. If $n = \ell^2$, the right side of the identity is

$$(-1)^{\ell-1}\ell f(\ell) = (-1)^{n-1}\ell^2 = (-1)^{n-1}n.$$

To obtain the left side of the identity, we recall the involution (13.1) on triples (u, d, δ) such that $u^2 + d\delta = n$ and δ is odd, and obtain

$$\sum_{\substack{u+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u u = 0.$$

Then

$$\begin{aligned}
\sum_{\substack{u+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u f(u+d) &= \sum_{\substack{u+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u (u+d) \\
&= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u u + \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u d \\
&= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^u d \\
&= \sum_{u^2<n} (-1)^u \sum_{\substack{n-u^2=d\delta \\ \delta\equiv 1 \pmod 2}} d \\
&= \sum_{|u|<\sqrt{n}} (-1)^u \sigma^*(n-u^2).
\end{aligned}$$

Therefore,

$$\sum_{|u|<\sqrt{n}} (-1)^u \sigma^*(n-u^2) = \{(-1)^{n-1}n\}_{n=\ell^2}.$$

This completes the proof.

Theorem 13.3 (Fermat) *An odd prime number p can be represented by the quadratic form x^2+y^2 if and only if $p \equiv 1 \pmod 4$.*

Proof. Since every square is congruent to 0 or 1 modulo 4, it follows that a sum of two squares must be congruent to 0, 1, or 2 modulo 4, and so no integer congruent to 3 modulo 4 can be represented as the sum of two squares.

Let p be an odd prime number. Then p is certainly not a square. By Lemma 13.3,

$$\sigma^*(p) = 2\sigma^*(p-1) - 2\sigma^*(p-4) + 2\sigma^*(p-9) - \cdots.$$

Since $\sigma^*(p) = p+1$, we have

$$\frac{p+1}{2} = \sigma^*(p-1^2) - \sigma^*(p-2^2) + \sigma^*(p-3^2) - \cdots.$$

If $p \equiv 1 \pmod 4$, then $(p+1)/2$ is an odd integer, and so at least one of the terms on right side of this equation must be odd. Thus, there exists a positive integer $b < \sqrt{n}$ such that $\sigma^*(p-b^2)$ is odd. By Lemma 13.2, $p - b^2 = a^2$ for some odd integer a. This completes the proof.

Theorem 13.4 *If p is a prime number such that $p \equiv 1 \pmod 4$, then there exist unique positive integers a and b such that a is odd, b is even, and $p = a^2 + b^2$.*

Proof. Let

$$p = a_1^2 + b_1^2 = a_2^2 + b_2^2,$$

where a_1 and a_2 are positive odd integers and b_1 and b_2 are positive even integers. We must prove that $a_1 = a_2$ and $b_1 = b_2$.

If $a_1 < a_2$, then $b_1 > b_2$ and there exist positive integers x and y such that

$$a_2 = a_1 + 2x$$

and

$$b_2 = b_1 - 2y.$$

Then

$$\begin{aligned} p &= a_2^2 + b_2^2 \\ &= (a_1+2x)^2 + (b_1-2y)^2 \\ &= a_1^2 + 4a_1x + 4x^2 + b_1^2 - 4b_1y + 4y^2 \\ &= p + 4a_1x + 4x^2 - 4b_1y + 4y^2, \end{aligned}$$

and so

$$x(a_1 + x) = y(b_1 - y).$$

Let $(x, y) = d$. Define the positive integers X and Y by $x = dX$ and $y = dY$. Then

$$X(a_1 + x) = Y(b_1 - y).$$

Since $(X, Y) = 1$, it follows that there exists a positive integer r such that

$$rY = a_1 + x = a_1 + dX$$

and

$$rX = b_1 - y = b_1 - dY.$$

Then $r^2 + d^2 \geq 2$ and $x^2 + y^2 \geq 2$, and

$$p = a_1^2 + b_1^2 = (rY - dX)^2 + (rX + dY)^2 = (r^2 + d^2)(X^2 + Y^2),$$

which is impossible, since p is prime and not composite. Therefore, $a_1 = a_2$ and $b_1 = b_2$, and the representation of a prime $p \equiv 1 \pmod 4$ as a sum of two squares is essentially unique.

Theorem 13.5 *An odd prime number p can be represented by the quadratic form $x^2 + 2y^2$ if and only if $p \equiv 1$ or $3 \pmod 8$.*

Proof. Since every square is congruent to 0, 1, or 4 modulo 8, it follows that an odd integer n is of the form $a^2 + 2b^2$ only if $n \equiv 1$ or $3 \pmod 8$.

Let a be a positive integer, $a < \sqrt{n}$. By Lemma 13.3, for every positive integer n we have

$$\sigma^*(n) = 2 \sum_{1 \leq u < \sqrt{n}} (-1)^{u-1} \sigma^*(n - u^2) + \{(-1)^{n-1} n\}_{n=\ell^2}. \tag{13.2}$$

Let $1 \leq u < \sqrt{n}$. Applying Lemma 13.3 to $n - u^2$, we have

$$\sigma^*(n-u^2) = 2 \sum_{1 \leq v^2 < n-u^2} (-1)^{v-1} \sigma^*(n-u^2-v^2) + \{(-1)^{n-u-1}(n-u^2)\}_{n-u^2=\ell_u^2}.$$

Inserting this into (13.2), we obtain

$$\begin{aligned} \sigma^*(n) &= 4 \sum_{\substack{u,v \geq 1 \\ u^2+v^2<n}} (-1)^{u+v} \sigma^*(n - u^2 - v^2) \\ &\quad + 2(-1)^n \sum_{1 \leq u < \sqrt{n}} \{n - u^2\}_{n-u^2=\ell_u^2} + \{(-1)^{n-1} n\}_{n=\ell^2}. \end{aligned}$$

If $u \neq v$ and $u^2 + v^2 = n$, then $v^2 + u^2 = n$ and the pairs (u, v) and (v, u) both appear in the first sum. Considering congruences modulo 8, we obtain

$$\begin{aligned} 4 \sum_{\substack{u,v \geq 1 \\ u^2+v^2<n}} (-1)^{u+v} \sigma^*(n - u^2 - v^2) &= 8 \sum_{\substack{1 \leq u < v \\ u^2+v^2<n}} (-1)^{u+v} \sigma^*(n - u^2 - v^2) + 4 \sum_{\substack{u \geq 1 \\ 2u^2<n}} \sigma^*(n - 2u^2) \\ &\equiv 4 \sum_{\substack{u \geq 1 \\ 2u^2<n}} \sigma^*(n - 2u^2) \pmod{8}. \end{aligned}$$

Therefore,

$$\begin{aligned} \sigma^*(n) \equiv & \; 4 \sum_{\substack{u \geq 1 \\ 2u^2<n}} \sigma^*(n - 2u^2) + 2(-1)^n \sum_{\substack{u \geq 1 \\ u^2<n}} \{n - u^2\}_{n-u^2=\ell_u^2} \\ & + \{(-1)^{n-1} n\}_{n=\ell^2} \pmod{8}. \end{aligned}$$

Let $p \equiv 3 \pmod{8}$. The prime number p is not a square, and, by Theorem 13.3, p is also not the sum of two squares. Therefore,

$$\{(-1)^{p-1} p\}_{p=\ell^2} = \{p - u^2\}_{p-u^2=\ell_u^2} = 0$$

for all u, and so

$$4 \sum_{\substack{u \geq 1 \\ 2u^2<n}} \sigma^*(n - 2u^2) \equiv \sigma^*(p) = p + 1 \equiv 4 \pmod{8}.$$

Dividing this congruence by 4, we obtain

$$\sum_{\substack{u \geq 1 \\ 2u^2<n}} \sigma^*(n - 2u^2) \equiv \frac{p+1}{4} \equiv 1 \pmod{2},$$

and so $\sigma^*(n - 2b^2)$ is odd for some integer b. Then $n - 2b^2 = a^2$ for some odd number a, and $n = a^2 + 2b^2$.

Let $p \equiv 1 \pmod{8}$. Then

$$\sigma^*(p) = p + 1 \equiv 2 \pmod{8}.$$

By Theorems 13.3 and 13.4, there exist unique positive integers a and b such that $p = a^2 + b^2$, where a is odd and b is even. This implies that

$$\sum_{\substack{u \geq 1 \\ u^2<p}} \{(p - u^2)\}_{p-u^2=\ell_u^2} = \{p - a^2\}_{p-a^2=b^2} + \{p - b^2\}_{p-b^2=a^2} = b^2 + a^2 = p,$$

and so

$$\begin{aligned}
2 &\equiv \sigma^*(p) \pmod 8 \\
&\equiv 4 \sum_{\substack{u \geq 1 \\ 2u^2 < p}} \sigma^*(p - 2u^2) + 2(-1)^p \sum_{1 \leq u^2 < p} \{(p - u^2)\}_{p-u^2=\ell_u^2} \pmod 8 \\
&\equiv 4 \sum_{\substack{u \geq 1 \\ 2u^2 < p}} \sigma^*(p - 2u^2) - 2p \pmod 8 \\
&\equiv 4 \sum_{\substack{u \geq 1 \\ 2u^2 < p}} \sigma^*(p - 2u^2) - 2 \pmod 8.
\end{aligned}$$

Therefore,

$$4 \sum_{\substack{u \geq 1 \\ 2u^2 < p}} \sigma^*(p - 2u^2) - 2 \equiv 2 \pmod 8,$$

and

$$\sum_{\substack{u \geq 1 \\ 2u^2 < p}} \sigma^*(p - 2u^2) \equiv 1 \pmod 2.$$

It follows that $\sigma^*(p-2b^2)$ is odd for some positive integer b, and so $p-2b^2 = a^2$ for some odd integer a. This completes the proof.

Exercises

1. Prove that $\sigma^*(n) = 1$ if and only if $n = 2^k$ for some nonnegative integer k.

2. Let $d(n)$ denote the number of positive divisors of n. Prove that $d(n)$ is odd if and only if n is a square.

3. Prove that n is a sum of two squares if and only if $2n$ is a sum of two squares. Hint: Consider the identity $2(x^2 + y^2) = (x + y)^2 + (x - y)^2$.

   Let $n = 2^k m$, where $k \geq 0$ and m is odd. Prove that n is a sum of two squares if and only if m is a sum of two squares.

4. Verify the polynomial identity

   $$(x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1x_2 - y_1y_2)^2 + (x_1y_2 + y_1x_2)^2.$$

   Deduce that if each of the integers n_1 and n_2 can be represented as a sum of two squares, then their product n_1n_2 is also a sum of two squares.

5. Let $k \geq 2$ and let $n_1, \ldots, n_k$ be positive integers. Prove that if each integer n_i is a sum of two squares, then the product $n_1n_2 \cdots n_k$ is a sum of two squares.

6. For every prime p and positive integer n, let $v_p(n)$ denote the highest power of p that divides n. Prove that if $v_p(n)$ is even for every prime $p \equiv 3 \pmod 4$, then n can be represented as a sum of two squares.

7. Let a and b be relatively prime integers, and let p be an odd prime. Prove that if p divides $a^2 + b^2$, then $p \equiv 1 \pmod 4$. Hint: Show that $(ab^{-1})^2 \equiv -1 \pmod p$, and so $\left(\frac{-1}{p}\right) = 1$, where $\left(\frac{a}{p}\right)$ is the Legendre symbol. Recall that $\left(\frac{-1}{p}\right) = 1$ if and only if $p \equiv 1 \pmod 4$.

8. Let p be a prime number, $p \equiv 3 \pmod 4$, and let a and b be integers. Prove that if p^c exactly divides a^2+b^2 (that is, p^c is the highest power of p that divides $a^2 + b^2$, then c is even. Hint: Let $d = (a, b)$, and let p^γ exactly divide d. Let $a = dA$ and $b = dB$, and consider the highest power of p that divides $A^2 + B^2$.

9. Prove that if n can be represented as a sum of two squares, then $v_p(n)$ is even for every prime $p \equiv 3 \pmod 4$.

13.3 A Ternary Form

We begin with the ternary quadratic form

$$Q(x, y, z) = x^2 + yz.$$

A *representation* of n by the quadratic form $Q(x, y, z)$ is an ordered triple of integers (x, y, z) such that $Q(x, y, z) = n$. We denote by $\mathcal{R}(n)$ the set of all representations of n by the quadratic form Q, that is,

$$\mathcal{R}(n) = \{(x, y, z) : Q(x, y, z) = n\}.$$

We introduce six bijections from the set $\mathcal{R}(n)$ to itself. The simplest are the involutions

$$\begin{aligned}
\rho(x, y, z) &= (x, z, y), \\
\sigma(x, y, z) &= (-x, y, z),
\end{aligned}$$

and

$$\tau(x, y, z) = (x, -y, -z).$$

Let

$$\alpha(x, y, z) = (z - x, 2x + y - z, z). \tag{13.3}$$

If $(x, y, z) \in \mathcal{R}(n)$, then

$$Q(\alpha(x, y, z)) = Q(z - x, 2x + y - z, z)$$

$$\begin{aligned} &= (z-x)^2 + (2x+y-z)z \\ &= z^2 - 2xz + x^2 + 2xz + yz - z^2 \\ &= x^2 + yz \\ &= n, \end{aligned}$$

and so $\alpha(x, y, z) \in \mathcal{R}(n)$. Moreover,

$$\alpha^2(x, y, z) = \alpha(z-x, 2x+y-z, z) = (x, y, z),$$

and so α is also an involution on the set $\mathcal{R}(n)$.

Let

$$\beta(x, y, z) = (x+y, y, -2x-y+z). \tag{13.4}$$

If $(x, y, z) \in \mathcal{R}(n)$, then

$$\begin{aligned} Q(\beta(x, y, z)) &= Q(x+y, y, -2x-y+z) \\ &= (x+y)^2 + y(-2x-y+z) \\ &= x^2 + 2xy + y^2 - 2xy - y^2 + yz \\ &= x^2 + yz \\ &= n, \end{aligned}$$

and so $\beta(x, y, z) \in \mathcal{R}(n)$.

Let

$$\gamma(x, y, z) = (x-y, y, 2x-y+z). \tag{13.5}$$

If $(x, y, z) \in \mathcal{R}(n)$, then

$$\begin{aligned} Q(\gamma(x, y, z)) &= Q(x-y, y, 2x-y+z) \\ &= (x-y)^2 + y(2x-y+z) \\ &= x^2 - 2xy + y^2 + 2xy - y^2 + yz \\ &= x^2 + yz \\ &= n, \end{aligned}$$

and so $\gamma(x, y, z) \in \mathcal{R}(n)$. Moreover,

$$\begin{aligned} \gamma\beta(x, y, z) &= \gamma(x+y, y, -2x-y+z) \\ &= (x+y-y, y, 2(x+y)-y+(-2x-y+z)) \\ &= (x, y, z). \end{aligned}$$

Similarly,

$$\beta\gamma(x, y, z) = (x, y, z).$$

Therefore, $\beta, \gamma : \mathcal{R}(n) \to \mathcal{R}(n)$ are bijections with $\gamma = \beta^{-1}$.

Finally, we state the following simple lemma, which will be used in the proof of Liouville's formula.

Lemma 13.4 *Let $\mathcal{S}$ and $\mathcal{S}'$ be finite sets, and let $\vartheta : \mathcal{S} \to \mathcal{S}'$ be a bijection with inverse $\vartheta^{-1} : \mathcal{S}' \to \mathcal{S}$. If $G(s)$ is a function defined for all $s \in \mathcal{S}$, then*

$$\sum_{s \in \mathcal{S}} G(s) = \sum_{s' \in \mathcal{S}'} G(\vartheta^{-1}(s')).$$

Proof. This follows instantly from the fact that $\vartheta^{-1}(\mathcal{S}') = \mathcal{S}$.

Exercises

1. Prove that $\sigma\beta\sigma = \gamma$ and $\rho\beta\sigma\rho = \alpha$.
2. Prove that $\beta\sigma$ is an involution.
3. Prove that

$$\beta^n(x, y, z) = (x + ny, y, z - 2nx - n^2 y).$$

4. Compute $\gamma^n(x, y, z)$.
5. Consider the 3×3 matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Let $\mathbf{v}$ denote the column vector

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Its transpose is $\mathbf{v}^T = (x, y, z)$. Show that

$$Q(x, y, z) = \mathbf{v}^T A \mathbf{v}.$$

6. Let $Q_1(x, y, z) = x^2 + y^2 - z^2$. Check that $Q(x, y + z, y - z) = Q_1(x, y, z)$ and $Q_1(x, (y + z)/2, (y - z)/2) = Q(x, y, z)$.

13.4 Proof of Liouville's Identity

In this section we prove Theorem 13.1.

For every positive integer n, we let $\mathcal{S}(n)$ be the set of all triples (u, d, δ) such that

$$Q(u, d, \delta) = u^2 + d\delta = n,$$

where u is an integer and d and δ are positive integers. Then $\mathcal{S}(n)$ is a finite subset of $\mathcal{R}(n)$. Using this notation, we have

$$\sum_{u^2+d\delta=n} = \sum_{(u,d,\delta)\in\mathcal{S}(n)}.$$

Partition $\mathcal{S}(n)$ into three sets $\mathcal{S}_1(n)$, $\mathcal{S}_{-1}(n)$, and $\mathcal{S}_0(n)$ as follows:

$$\mathcal{S}_1(n) = \{(u,d,\delta) \in \mathcal{S}(n) : 2u + d - \delta \geq 1\},$$

$$\mathcal{S}_0(n) = \{(u,d,\delta) \in \mathcal{S}(n) : 2u + d - \delta = 0\},$$

and

$$\mathcal{S}_{-1}(n) = \{(u,d,\delta) \in \mathcal{S}(n) : 2u + d - \delta \leq -1\}.$$

Let α be the map on $\mathcal{S}(n)$ defined by (13.3). If $(u,d,\delta) \in \mathcal{S}(n)$, then d and δ are positive integers. If $(u,d,\delta) \in \mathcal{S}_1(n)$, then $2u + d - \delta \geq 1$, and so

$$(u',d',\delta') = \alpha(u,d,\delta) = (\delta - u, 2u + d - \delta, \delta) \in \mathcal{S}(n).$$

Since

$$2u' + d' - \delta' = 2(\delta - u) + (2u + d - \delta) - \delta = d \geq 1,$$

it follows that $\alpha(u,d,\delta) \in \mathcal{S}_1(n)$, and so α is an involution on $\mathcal{S}_1(n)$. Moreover,

$$\begin{aligned} \delta' - 2u' &= \delta - 2(\delta - u) = -(\delta - 2u), \\ u' + d' &= (\delta - u) + (2u + d - \delta) = u + d, \end{aligned}$$

and

$$2u' + 2d' - \delta' = 2(\delta - u) + 2(2u + d - \delta) - \delta = 2u + 2d - \delta.$$

Let $F(x,y,z)$ be a function that is odd in x and even in the pair (y,z). We define the function

$$G(x,y,z) = F(z - 2x, x + y, 2x + 2y - z).$$

If $(u,d,\delta) \in \mathcal{S}_1(n)$ and $\alpha(u,d,\delta) = (u',d',\delta')$, then

$$\begin{aligned} & G(u,d,\delta) + G(u',d',\delta') \\ &= F(\delta - 2u, u + d, 2u + 2d - \delta) + F(\delta' - 2u', u' + d', 2u' + 2d' - \delta') \\ &= F(\delta - 2u, u + d, 2u + 2d - \delta) + F(-(\delta - 2u), u + d, 2u + 2d - \delta) \\ &= 0, \end{aligned}$$

since the function $F(x,y,z)$ is odd in its first variable x. From Lemma 13.4 with $\mathcal{S} = \mathcal{S}' = \mathcal{S}_1(n)$ and $\vartheta = \vartheta^{-1} = \alpha$, we obtain

$$\begin{aligned}\sum_{(u,d,\delta)\in\mathcal{S}_1(n)} F(\delta-2u, u+d, 2u+2d-\delta) &= \sum_{(u,d,\delta)\in\mathcal{S}_1(n)} G(u,d,\delta)\\ &= \sum_{(u,d,\delta)\in\mathcal{S}_1(n)} G(u',d',\delta')\\ &= -\sum_{(u,d,\delta)\in\mathcal{S}_1(n)} G(u,d,\delta)\\ &= 0.\end{aligned}$$

Next we consider triples $(u,d,\delta)\in\mathcal{S}_0(n)$. Since

$$2u+d-\delta = 0,$$

it follows that

$$u = \frac{\delta-d}{2}$$

and

$$n = u^2+d\delta = \left(\frac{\delta-d}{2}\right)^2 + d\delta = \left(\frac{d+\delta}{2}\right)^2 = \ell^2,$$

where

$$\ell = \frac{d+\delta}{2} \geq 1.$$

Therefore, the set $\mathcal{S}_0(n)$ is nonempty only if n is a square. Moreover, the integers d and δ are positive, and so

$$1 \leq d = 2\ell-\delta \leq 2\ell-1.$$

Conversely, if $1\leq d\leq 2\ell-1$, we set $\delta = 2\ell-d$ and $u=\ell-d$. Then

$$\begin{aligned}u^2+d\delta &= (\ell-d)^2+d(2\ell-d) = \ell^2 = n,\\ 2u+d-\delta &= 0,\end{aligned}$$

and

$$(u,d,\delta)\in\mathcal{S}_0(n).$$

It follows that if $n=\ell^2$ with $\ell\geq 1$, then

$$\mathcal{S}_0(n) = \{(d-\ell, d, 2\ell-d) : 1\leq d\leq 2\ell-1\}$$

and

$$\sum_{(u,d,\delta)\in\mathcal{S}_0(n)} F(\delta-2u, u+d, 2u+2d-\delta) = \sum_{d=1}^{2\ell-1} F(d,\ell,d) = T_1(n).$$

To analyze the sum

$$\sum_{(u,d,\delta)\in\mathcal{S}(n)} F(d+\delta,u,d-\delta),$$

we construct a second partition of $\mathcal{S}(n)$. Define the three sets $\mathcal{S}'_1(n)$, $\mathcal{S}'_{-1}(n)$, and $\mathcal{S}'_0(n)(n)$ as follows:

$$\begin{aligned}\mathcal{S}'_1(n) &= \{(u,d,\delta)\in\mathcal{S}(n) : 2u-d+\delta\geq 1\},\\ \mathcal{S}'_{-1}(n) &= \{(u,d,\delta)\in\mathcal{S}(n) : 2u-d+\delta\leq -1\},\end{aligned}$$

and

$$\mathcal{S}'_0(n) = \{(u,d,\delta)\in\mathcal{S}(n) : 2u-d+\delta = 0\}.$$

We shall prove that

$$\begin{aligned}&\sum_{(u,d,\delta)\in\mathcal{S}_{-1}(n)} F(\delta-2u,u+d,2u+2d-\delta)\\ &= \sum_{(u,d,\delta)\in\mathcal{S}'_1(n)} F(d+\delta,u,d-\delta)\\ &= \sum_{(u,d,\delta)\in\mathcal{S}'_{-1}(n)} F(d+\delta,u,d-\delta)\end{aligned}$$

and

$$\sum_{(u,d,\delta)\in\mathcal{S}'_0(n)} F(d+\delta,u,d-\delta) = \{T_2(n)\}_{n=\ell^2}.$$

Let β be the map on $\mathcal{S}(n)$ defined by (13.4). If $(u,d,\delta)\in\mathcal{S}_{-1}(n)$, then $2u+d-\delta\leq -1$, and so $-2u-d+\delta\geq 1$ and

$$(u',d',\delta') = \beta(u,d,\delta) = (u+d,d,-2u-d+\delta)\in\mathcal{S}(n).$$

Moreover,

$$2u'-d'+\delta' = 2(u+d)-d+(-2u-d+\delta) = \delta\geq 1,$$

and so

$$\beta : \mathcal{S}_{-1}(n)\to\mathcal{S}'_1(n).$$

Let γ be the map on $\mathcal{S}(n)$ defined by (13.5). If $(u',d',\delta')\in\mathcal{S}'_1(n)$, then $2u'-d'+\delta'\geq 1$ and

$$(u,d,\delta) = \gamma(u',d',\delta') = (u'-d',d',2u'-d'+\delta')\in\mathcal{S}(n).$$

Moreover,

$$2u+d-\delta = 2(u'-d')+d'-(2u'-d'+\delta') = -\delta'\leq -1,$$

and so $(u,d,\delta) \in \mathcal{S}_{-1}(n)$. Therefore, the map

$$\gamma : \mathcal{S}'_1(n) \to \mathcal{S}_{-1}(n)$$

is a bijection, and $\gamma = \beta^{-1}$.

Applying Lemma 13.4, we obtain

$$\begin{aligned}
&\sum_{(u,d,\delta)\in\mathcal{S}_{-1}(n)} F(\delta - 2u, u + d, 2u + 2d - \delta) \\
&= \sum_{(u,d,\delta)\in\mathcal{S}_{-1}(n)} G(u,d,\delta) \\
&= \sum_{(u',d',\delta')\in\beta(\mathcal{S}_{-1}(n))} G(\gamma(u',d',\delta')) \\
&= \sum_{(u',d',\delta')\in\mathcal{S}'_1(n)} G(u'-d', d', 2u'-d'+\delta') \\
&= \sum_{(u',d',\delta')\in\mathcal{S}'_1(n)} F(d'+\delta', u', d'-\delta').
\end{aligned}$$

Let ψ be the map on $\mathcal{S}(n)$ defined by $\psi(u,d,\delta) = (-u,\delta,d)$. Then ψ is an involution since $\psi = \rho\sigma$. If $(u,d,\delta) \in \mathcal{S}'_1(n)$, then $2u - d + \delta \geq 1$, and so $-2u - \delta + d \leq -1$ and

$$\psi(u,d,\delta) = (-u,\delta,d) \in \mathcal{S}'_{-1}(n).$$

Similarly, if $(u,d,\delta) \in \mathcal{S}'_{-1}(n)$, then $2u - d + \delta \geq 1$, and so $-2u - \delta + d \geq 1$ and

$$\psi(u,d,\delta) = (-u,\delta,d) \in \mathcal{S}'_1(n).$$

Therefore,

$$\psi : \mathcal{S}'_1(n) \to \mathcal{S}'_{-1}(n)$$

is a bijection with $\psi^{-1} = \psi$. Let

$$H(x,y,z) = F(y+z, x, y-z).$$

By Lemma 13.4,

$$\begin{aligned}
\sum_{(u,d,\delta)\in\mathcal{S}'_1(n)} F(d+\delta, u, d-\delta) &= \sum_{(u,d,\delta)\in\mathcal{S}'_1(n)} H(u,d,\delta) \\
&= \sum_{(u,d,\delta)\in\mathcal{S}'_{-1}(n)} H(\psi(u,d,\delta)) \\
&= \sum_{(u,d,\delta)\in\mathcal{S}'_{-1}(n)} H(-u,\delta,d) \\
&= \sum_{(u,d,\delta)\in\mathcal{S}'_{-1}(n)} F(\delta+d, -u, -\delta-d) \\
&= \sum_{(u,d,\delta)\in\mathcal{S}'_{-1}(n)} F(d+\delta, u, d+\delta),
\end{aligned}$$

since the function $F(x,y,z)$ is even in the pair of variables (y,z).

If $(u,d,\delta) \in \mathcal{S}_0'(n)$, then

$$\begin{aligned} 2u - d + \delta &= 0, \\ u &= \frac{d-\delta}{2}, \end{aligned}$$

and

$$n = u^2 + d\delta = \left(\frac{d-\delta}{2}\right)^2 + d\delta = \left(\frac{d+\delta}{2}\right)^2 = \ell^2,$$

where

$$\ell = \frac{d+\delta}{2}.$$

Therefore, the set $\mathcal{S}_0'(n)$ is nonempty only if n is a square. Since the integers d and δ are positive, it follows that

$$1 \leq d = 2\ell - \delta \leq 2\ell - 1.$$

Conversely, if $1 \leq d \leq 2\ell - 1$, we set $\delta = 2\ell - d$ and $u = d - \ell$. Then

$$\begin{aligned} u^2 + d\delta &= (d-\ell)^2 + d(2\ell - d) = \ell^2 = n, \\ 2u - d + \delta &= 0, \end{aligned}$$

and

$$(u,d,\delta) \in \mathcal{S}_0'(n).$$

It follows that if $n = \ell^2$ with $\ell \geq 1$, then

$$\mathcal{S}_0'(n) = \{(d-\ell, d, 2\ell - d) : 1 \leq d \leq 2\ell - 1\}$$

and

$$\begin{aligned} \sum_{(u,d,\delta)\in\mathcal{S}_0'(n)} F(d+\delta, u, d-\delta) &= \sum_{d=1}^{2\ell-1} F(2\ell, d-\ell, 2d-2\ell) \\ &= \sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j) \\ &= T_2(n). \end{aligned}$$

Therefore,

$$\begin{aligned} &\sum_{(u,d,\delta)\in\mathcal{S}(n)} F(d+\delta, u, d-\delta) \\ &= 2 \sum_{(u,d,\delta)\in\mathcal{S}_1'(n)} F(d+\delta, u, d-\delta) + \{T_2(n)\}_{n=\ell^2} \end{aligned}$$

$$
\begin{aligned}
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}_{-1}(n)} F(\delta-2u, u+d, 2u+2d-\delta) + \{T_2(n)\}_{n=\ell^2} \\
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}_{-1}(n)} F(\delta-2u, u+d, 2u+2d-\delta) + \\
&\quad 2 \sum_{(u,d,\delta)\in\mathcal{S}_{1}(n)} F(\delta-2u, u+d, 2u+2d-\delta) + \{T_2(n)\}_{n=\ell^2} \\
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}(n)} F(\delta-2u, u+d, 2u+2d-\delta) \\
&\quad - 2\{T_1(n)\}_{n=\ell^2} + \{T_2(n)\}_{n=\ell^2}.
\end{aligned}
$$

This completes the proof of Theorem 13.1.

13.5 Two Corollaries

In this section we derive two additional identities that we use in the next chapter.

Theorem 13.6 *If $F(x,y,z)$ is a function that is odd in each of the variables x, y, and z, and if $F(x,y,z) = 0$ for every even integer x, then*

$$
\sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ \delta\equiv 1 \pmod 2}} F(\delta-2u, u+d, 2u+2d-\delta) = \{T_0(\ell)\}_{n=\ell^2},
$$

where

$$
T_0(\ell) = \sum_{j=1}^{\ell} F(2j-1, \ell, 2j-1).
$$

Proof. Since the function $F(x,y,z)$ is odd in the variable y, we have $F(x,0,z) = 0$ for all x and z, and

$$
\begin{aligned}
&\sum_{(u,d,\delta)\in\mathcal{S}(n)} F(d+\delta, u, d-\delta) \\
&= \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\geq 1}} F(d+\delta, u, d-\delta) + \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\leq -1}} F(d+\delta, u, d-\delta) \\
&= \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\geq 1}} F(d+\delta, u, d-\delta) + \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\geq 1}} F(d+\delta, -u, d-\delta) \\
&= \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\geq 1}} F(d+\delta, u, d-\delta) - \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n) \\ u\geq 1}} F(d+\delta, u, d-\delta) \\
&= 0.
\end{aligned}
$$

Since $F(x,y,z)=0$ for all even integers x, we have

$$\sum_{(u,d,\delta)\in\mathcal{S}(n)} F(\delta-2u,u+d,2u+2d-\delta) = \sum_{\substack{(u,d,\delta)\in\mathcal{S}(n)\\ \delta\equiv 1 \pmod 2}} F(\delta-2u,u+d,2u+2d-\delta).$$

If $n=\ell^2$, then

$$T_1(\ell)=\sum_{j=1}^{2\ell-1}F(j,\ell,j)=\sum_{j=1}^{\ell}F(2j-1,\ell,2j-1)$$

and

$$T_2(\ell)=\sum_{j=-\ell+1}^{\ell-1}F(2\ell,j,2j)=0.$$

The result follows immediately from Theorem 13.1.

Theorem 13.7 *Let $f(x,y)$ be a function that is odd in each of the variables x and y. For every positive integer n,*

$$\sum_{\substack{u^2+d\delta=n\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}f(\delta-2u,u+d)=\{T_0(\ell)\}_{n=\ell^2},$$

where

$$T_0(\ell)=\sum_{j=1}^{\ell}(-1)^{j+\ell}f(2j-1,\ell).$$

Proof. We define the function $F(x,y,z)$ as follows:

$$F(x,y,z)=\begin{cases}0 & \text{if } x \text{ or } z \text{ is even,}\\ (-1)^{y+\frac{z+1}{2}}f(x,y) & \text{if } x \text{ and } z \text{ are odd.}\end{cases}$$

Then $F(x,y,z)$ is a function that is odd in each of the variables x,y, and z, and $F(x,y,z)=0$ for every even integer x. By Theorem 13.6, we have

$$\begin{aligned}&\sum_{\substack{u^2+d\delta=n\\ \delta\equiv 1 \pmod 2}}F(\delta-2u,u+d,2u+2d-\delta)\\ &= \sum_{\substack{u^2+d\delta=n\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}f(\delta-2u,u+d)\\ &= \{T_0(\ell)\}_{n=\ell^2},\end{aligned}$$

where

$$\begin{aligned}T_0(\ell) &= \sum_{j=1}^{\ell}F(2j-1,\ell,2j-1)\\ &= \sum_{j=1}^{\ell}(-1)^{j+\ell}f(2j-1,\ell).\end{aligned}$$

This completes the proof.

13.6 Notes

Liouville's papers contain the statements of many theorems, but no proofs. Dickson's *History of the Theory of Numbers* [25], Volume II, Chapter XI, "Liouville's series of eighteen articles," contains a detailed summary of Liouville's assertions and references to papers by other mathematicians who have provided proofs of Liouville's results.

Uspensky and Heaslet [145] and Venkov [149] present careful accounts of Liouville's method and proofs of many of his results.

14

Sums of an Even Number of Squares

> The problem of the representation of an integer n as the sum of a given number k of integral squares is one of the most celebrated in the theory of numbers.... Almost every arithmetician of note since Fermat has contributed to the solution of the problem, and it has its puzzles for us still.
>
> G. H. Hardy [52, p. 132]

14.1 Summary of Results

For every positive integer s and nonnegative integer n, we let $R_s(n)$ denote the number of ordered s-tuples of integers $(x_1, \ldots, x_s)$ such that

$$n = x_1^2 + \cdots + x_s^2.$$

The integers x_i can be positive, negative, or 0. For every $s \geq 1$ we have

$$R_s(0) = 1,$$

since $0 = 0^2 + \cdots + 0^2$ is the unique representation of 0 as a sum of squares.

We shall apply Liouville's identities to obtain explicit formulae for the number of representations of a positive integer as the sum of s squares, where $s = 2, 4, 6, 8$, and 10. Representing an integer n as the sum of s squares is a problem in additive number theory, but the solution, for even

values of s, always involves a sum over the *divisors* of n, a fundamental topic in multiplicative number theory.

In this chapter, d and δ always denote positive integers, and $\sum_{d|n}$ and $\sum_{n=d\delta}$ denote the sum over the positive divisors of n.

We write the positive integer n in the form $n = 2^a m$, where $a \geq 0$ and m is odd. We shall prove the following formulae:

$$
\begin{aligned}
R_2(n) &= 4\sum_{d|m}(-1)^{(d-1)/2}, \\
R_4(n) &= \begin{cases} 8\sum_{d|n} d & \text{if } n \text{ is odd,} \\ 24\sum_{d|m} d & \text{if } n \text{ is even,} \end{cases} \\
R_6(n) &= 4\left(4^{a+1} - (-1)^{(m-1)/2}\right) \sum_{m=d\delta} (-1)^{(\delta-1)/2} d^2, \\
R_8(n) &= \begin{cases} 16\sum_{d|n} d^3 & \text{if } n \text{ is odd,} \\ (16/7)(8^{a+1} - 15)\sum_{d|m} d^3 & \text{if } n \text{ is even,} \end{cases} \\
R_{10}(n) &= \frac{4}{5}\left(16^{a+1} + (-1)^{(m-1)/2}\right) \sum_{m=d\delta} (-1)^{(\delta-1)/2} d^4 \\
&\quad + \frac{16}{5} \sum_{n=v^2+w^2} \left(v^4 - 3v^2w^2\right).
\end{aligned}
$$

14.2 A Recursion Formula

Our proofs depend on the following recursion formula for $R_s(n)$.

Theorem 14.1 *For all positive integers s and n,*

$$\sum_{|u|\leq\sqrt{n}} \left(n - (s+1)u^2\right) R_s(n-u^2) = 0. \tag{14.1}$$

Proof. If

$$n = x_1^2 + \cdots + x_s^2 + x_{s+1}^2,$$

then $x_{s+1}^2 \leq n$ and so

$$|x_{s+1}| \leq \sqrt{n}.$$

For $j = 1, \ldots, R_{s+1}(n)$, let

$$n = \sum_{i=1}^{s+1} x_{i,j}^2$$

denote the $R_{s+1}(n)$ representations of n as a sum of $s+1$ squares. For $i = 1, \ldots, s$, we define the map τ_i on the set of $(s+1)$-tuples by

$$\tau_i(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_s, x_{s+1}) = (x_1, \ldots, x_{i-1}, x_{s+1}, x_{i+1}, \ldots, x_s, x_i).$$

This is an involution on the set of the $R_{s+1}(n)$ representations of n as a sum of $s+1$ squares, and so

$$\sum_{j=1}^{R_{s+1}(n)} x_{s+1,j}^2 = \sum_{j=1}^{R_{s+1}(n)} x_{i,j}^2 \qquad \text{for } i = 1, \ldots, s.$$

Summing over all representations of n, we obtain

$$\begin{aligned} nR_{s+1}(n) &= \sum_{j=1}^{R_{s+1}(n)} \sum_{i=1}^{s+1} x_{i,j}^2 \\ &= \sum_{i=1}^{s+1} \sum_{j=1}^{R_{s+1}(n)} x_{i,j}^2 \\ &= (s+1) \sum_{j=1}^{R_{s+1}(n)} x_{s+1,j}^2 \\ &= (s+1) \sum_{|u|\leq\sqrt{n}} u^2 R_s(n-u^2), \end{aligned}$$

since for every integer u with $|u| \leq \sqrt{n}$ there are $R_s(n-u^2)$ representations $n = \sum_{i=1}^{s+1} x_{i,j}^2$ with $x_{s+1,j} = u$. This also implies that

$$R_{s+1}(n) = \sum_{|u|\leq\sqrt{n}} R_s(n-u^2).$$

Then

$$nR_{s+1}(n) = n \sum_{|u|\leq\sqrt{n}} R_s(n-u^2),$$

and

$$\sum_{|u|\leq\sqrt{n}} \left(n - (s+1)u^2\right) R_s(n-u^2) = 0.$$

This completes the proof.

Theorem 14.2 *Let $\Phi(n)$ be a function defined for all nonnegative integers n such that*

$$\Phi(0) = 1$$

and

$$\sum_{|u|\leq\sqrt{n}} \left(n - (s+1)u^2\right) \Phi(n-u^2) = 0$$

for $n \geq 1$. Then

$$\Phi(n) = R_s(n)$$

for all $n \geq 0$.

Proof. This follows immediately from Theorem 14.1.

The recursion formula (14.1) enables us to compute $R_s(n)$ for all positive integers s and n. We have

$$\begin{aligned} nR_s(n) &= -\sum_{1\leq |u|\leq\sqrt{n}} \left(n-(s+1)u^2\right) R_s(n-u^2) \\ &= 2\sum_{1\leq u\leq\sqrt{n}} \left((s+1)u^2-n\right) R_s(n-u^2), \end{aligned}$$

and so

$$R_s(n) = 2\sum_{1\leq u\leq\sqrt{n}} \left(\frac{(s+1)u^2}{n}-1\right) R_s(n-u^2). \tag{14.2}$$

For example, for $s=3$ we have

$$\begin{aligned}
R_3(1) &= 2\left(\tfrac{4\cdot 1^2}{1}-1\right) R_3(1-1^2) &= 6\ ,\\
R_3(2) &= 2\left(\tfrac{4\cdot 1^2}{2}-1\right) R_3(2-1^2) &= 12\ ,\\
R_3(3) &= 2\left(\tfrac{4\cdot 1^2}{3}-1\right) R_3(3-1^2) &= 8\ ,\\
R_3(4) &= 2\left(\left(\tfrac{4\cdot 1^2}{4}-1\right) R_3(4-1^2)+\left(\tfrac{4\cdot 2^2}{4}-1\right) R_3(4-2^2)\right) &= 6\ ,\\
R_3(5) &= 2\left(\left(\tfrac{4\cdot 1^2}{5}-1\right) R_3(5-1^2)+\left(\tfrac{4\cdot 2^2}{5}-1\right) R_3(5-2^2)\right) &= 24\ ,\\
R_3(6) &= 2\left(\left(\tfrac{4\cdot 1^2}{6}-1\right) R_3(6-1^2)+\left(\tfrac{4\cdot 2^2}{6}-1\right) R_3(6-2^2)\right) &= 24\ ,\\
R_3(7) &= 2\left(\left(\tfrac{4\cdot 1^2}{7}-1\right) R_3(7-1^2)+\left(\tfrac{4\cdot 2^2}{7}-1\right) R_3(7-2^2)\right) &= 0\ ,\\
R_3(8) &= 2\left(\left(\tfrac{4\cdot 1^2}{8}-1\right) R_3(8-1^2)+\left(\tfrac{4\cdot 2^2}{8}-1\right) R_3(8-2^2)\right) &= 12\ .
\end{aligned}$$

Exercises

1. Prove that $R_s(n) < R_{s+1}(n)$ for all positive integers s and n.
2. Use induction on s to prove (without using Theorem 14.1) that $R_s(n)$ is even for all positive integers s and n.
3. Use the recursion formula (14.2) to compute $R_2(n)$ and $R_4(n)$ for $n \leq 8$.
4. For positive integers k and s, let $R_{k,s}(n)$ denote the number of s-tuples of integers such that
$$x_1^k+\cdots+x_s^k = n.$$
Then $R_s(n) = R_{2,s}(n)$ and $R_{2k,s}(0) = 1$. Prove that
$$\sum_{|u|\leq n^{1/2k}} \left(n-(s+1)u^{2k}\right) R_{2k,s}(n-u^{2k}) = 0$$
for every positive integer n.

5. Let k and s be positive integers. Prove that

$$R_{2k,s}(1) = 2s.$$

6. Let k and s be positive integers, and let $0 \leq n < 4^k$. Prove that

$$R_{2k,s}(n) = 2^n \binom{s}{n}.$$

7. Let $s \geq 3$. Show that $R_{3,s}(n^3) = \infty$ for every integer n.

8. For positive integers k and s, let $r_{k,s}(n)$ denote the number of s-tuples of nonnegative integers such that

$$x_1^k + \cdots + x_s^k = n.$$

Prove that $r_{k,s}(0) = 1$ and

$$\sum_{0 \leq u \leq n^{1/k}} \left(n - (s+1)u^k\right) r_{k,s}(n - u^k) = 0$$

for every positive integer n.

14.3 Sums of Two Squares

Recall that $\mathcal{S}(n)$ is the set of all triples (u, d, δ) of integers with $d, \delta \geq 1$ and $u^2 + d\delta = n$.

If k_1 and k_2 are odd integers, then the function $f(x, y) = x^{k_1} y^{k_2}$ is odd in each of the variables x and y. Applying Theorem 13.7, we obtain

$$\sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} (\delta - 2u)^{k_1} (d+u)^{k_2} = \left\{ \ell^{k_2} \sum_{j=1}^{\ell} (-1)^{\ell-j} (2j-1)^{k_1} \right\}_{n=\ell^2}. \quad (14.3)$$

We shall use this identity for various values of k_1 and k_2. We can simplify the sum on the left by noticing that $(u, d, \delta) \in \mathcal{S}(n)$ if and only if $(-u, d, \delta) \in \mathcal{S}(n)$. This implies that if k is an odd integer and $g(d, \delta)$ is any function, then

$$\sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} u^k g(d, \delta) = 0. \quad (14.4)$$

Since $(u,d,\delta) \in \mathcal{S}(n)$ if and only if $(u,\delta,d) \in \mathcal{S}(n)$, it also follows that if $\varepsilon(d,\delta) = \varepsilon(\delta,d)$, then

$$\sum_{u^2+d\delta=n} \varepsilon(d,\delta)(d-\delta)h(u) = 0 \tag{14.5}$$

for any function $h(u)$.

In this section we shall obtain a formula for the number of representations of an integer as the sum of two squares. By Theorem 14.2, it suffices to construct a function $\Phi(n)$ such that $\Phi(0) = 1$ and

$$\sum_{|x|\leq\sqrt{n}} \left(n - 3x^2\right) \Phi(n - x^2) = 0$$

for every positive integer n.

Theorem 14.3

$$R_2(n) = 4\sum_{d|n}(-1)^{(d-1)/2} = 4\left(\sum_{\substack{d|n \\ d\equiv 1 \pmod 4}} 1 - \sum_{\substack{d|n \\ d\equiv 3 \pmod 4}} 1\right).$$

Proof. The function $f(x,y) = xy$ is odd in each of the variables x and y. The left side of identity (14.3) is

$$\begin{aligned}
&\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} f(\delta - 2u, d+u) \\
= &\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} (\delta - 2u)(d+u) \\
= &\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} (d\delta - 2u^2 + \delta u - 2du) \\
= &\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} (d\delta - 2u^2),
\end{aligned}$$

by (14.4) with $k = 1$.

If $n = \ell^2$, then (by Exercise 1) the right side of the identity (14.3) is

$$\begin{aligned}
T_0(\ell) &= \ell\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1) \\
&= \ell^2 \\
&= n.
\end{aligned}$$

Therefore,

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d\delta - 2u^2) = \{T_0(\ell)\}_{n=\ell^2}.$$

If d and δ are positive integers and

$$n = u^2 + d\delta,$$

then

$$|u| < \sqrt{n}$$

and

$$d\delta - 2u^2 = n - 3u^2.$$

Therefore,

$$\begin{aligned}\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d\delta - 2u^2) &= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(n - 3u^2) \\ &= \sum_{|u|<n} (n-3u^2) \sum_{\substack{\delta\mid(n-u^2) \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}.\end{aligned}$$

Define the function $\Phi(n)$ by $\Phi(0) = 1$ and, for every positive integer n,

$$\Phi(n) = 4 \sum_{\substack{\delta\mid n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}.$$

Then

$$\sum_{|u|<\sqrt{n}} (n-3u^2)\Phi(n-u^2) = \{4n\}_{n=\ell^2}.$$

If n is not a square, then

$$\sum_{|u|\leq\sqrt{n}} (n-3u^2)\Phi(n-u^2) = \sum_{|u|<\sqrt{n}} (n-3u^2)\Phi(n-u^2) = \{4n\}_{n=\ell^2} = 0.$$

If $n = \ell^2$ is a square, then

$$\begin{aligned}\sum_{|u|\leq\sqrt{n}} (n-3u^2)\Phi(n-u^2) &= \sum_{|u|<\sqrt{n}} (n-3u^2)\Phi(n-u^2) \\ &\quad +(n-3m^2)\Phi(0) + (n-3(-m)^2)\Phi(0) \\ &= \{4n\}_{n=\ell^2} - 2n - 2n \\ &= 0.\end{aligned}$$

Therefore,

$$R_2(n) = \Phi(n)$$

for all positive integers n. This completes the proof.

Exercises

1. Prove that for every positive integer ℓ,
$$\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)=\ell.$$

2. Let p be a prime number such that $p \equiv 1 \pmod{4}$. Prove that
$$R_2(p^k)=4(k+1).$$

3. Let p be a prime number such that $p \equiv 3 \pmod{4}$. Prove that
$$R_2(p^k)=\begin{cases} 4 & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.}\end{cases}$$

4. Define the divisor functions
$$d_1(n)=\sum_{\substack{d\mid n \\ d\equiv 1 \pmod{4}}} 1$$
and
$$d_3(n)=\sum_{\substack{d\mid n \\ d\equiv 3 \pmod{4}}} 1.$$
Prove that $d_1(n) \geq d_3(n)$ for every positive integer n.

5. Let p be a prime number, $p \equiv 3 \pmod{4}$. Prove that if $n = p^{2k-1}m$, where $(p,m)=1$, then
$$d_1(n)=kd_1(m)+kd_3(m)$$
and
$$d_3(n)=kd_1(m)+kd_3(m).$$
Deduce that n cannot be written as the sum of two squares.

6. An arithmetic function $f(n)$ is called *multiplicative* if
$$f(n_1n_2)=f(n_1)f(n_2)$$
for all positive integers n_1 and n_2 such that $(n_1,n_2)=1$. Define the function $\chi(n)$ by
$$\chi(n)=\begin{cases} 0 & \text{if } n \text{ is even,} \\ 1 & \text{if } n \equiv 1 \pmod{4}, \\ -1 & \text{if } n\equiv 3 \pmod{4}.\end{cases}$$

Prove that $\chi(n)$ is multiplicative.

Prove that

$$R_2(n) = \sum_{d|n} \chi(n).$$

Prove that $R_2(n)$ is multiplicative.

Hint: If $(n_1, n_2) = 1$ and d is a divisor of $n_1 n_2$, then there exist unique divisors d_1 of n_1 and d_2 of n_2 such that $d = d_1 d_2$.

7. The *divisor function* counts the number of positive divisors of n, that is,

$$d(n) = \sum_{d|n} 1.$$

Prove that $d(n)$ is a multiplicative function, and that

$$R_2(n) \leq 4d(n)$$

for all positive integers n.

Hint: Since $R_2(n)$ and $d(n)$ are both multiplicative functions, it suffices to to check the inequality for prime powers.

8. Prove that $\liminf_{n\to\infty} R_2(n) = 0$.

9. Prove that $\limsup_{n\to\infty} R_2(n) = \infty$.

14.4 Sums of Four Squares

In this section we prove Jacobi's formula for the number of representations of an integer as the sum of four squares.

Theorem 14.4 (Jacobi) *For every positive integer n,*

$$R_4(n) = 8\sum_{d|n} d \qquad \text{if } n \text{ is odd,}$$

and

$$R_4(n) = 24 \sum_{\substack{d|n \\ d\equiv 1 \pmod 2}} d \qquad \text{if } n \text{ is even.}$$

Proof. By Theorem 13.1, if $F(x, y, z)$ is a function of integer variables x, y, z that is odd in x and even in the pair (y, z), then

$$2\sum_{u^2+d\delta=n} F(\delta - 2u, u + d, 2u + 2d - \delta) - \sum_{u^2+d\delta=n} F(d + \delta, u, d - \delta)$$
$$= \left\{ 2\sum_{j=1}^{2\ell-1} F(j, \ell, j) - \sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j) \right\}_{n=\ell^2}.$$

The function $(-1)^x F(x,y,z)$ is also odd in x and even in the pair (y,z). Applying Theorem 13.1 to the function $(-1)^x F(x,y,z)$, we obtain

$$\begin{aligned} 2 \sum_{u^2+d\delta=n} & (-1)^\delta F(\delta-2u, u+d, 2u+2d-\delta) \\ & - \sum_{u^2+d\delta=n} (-1)^{d+\delta} F(d+\delta, u, d-\delta) \\ = & \left\{ 2\sum_{j=1}^{2\ell-1} (-1)^j F(j,\ell,j) - \sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j) \right\}_{n=\ell^2}. \end{aligned}$$

Adding these identities gives

$$\begin{aligned} 4 \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 0 \pmod 2}} & F(\delta-2u, u+d, 2u+2d-\delta) \\ & - 2 \sum_{\substack{u^2+d\delta=n \\ d\equiv \delta \pmod 2}} F(d+\delta, u, d-\delta) \\ = & \left\{ 4 \sum_{\substack{1\le j\le 2\ell-1 \\ j\equiv 0 \pmod 2}} F(j,\ell,j) - 2\sum_{j=-\ell+1}^{\ell-1} F(2\ell, j, 2j) \right\}_{n=\ell^2}. \qquad (14.6) \end{aligned}$$

Subtracting these identities gives

$$\begin{aligned} 4 \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} & F(\delta-2u, u+d, 2u+2d-\delta) \\ & - 2 \sum_{\substack{u^2+d\delta=n \\ d\equiv -\delta \pmod 2}} F(d+\delta, u, d-\delta) \\ = & \left\{ 4 \sum_{\substack{1\le j\le 2\ell-1 \\ j\equiv 1 \pmod 2}} F(j,\ell,j) \right\}_{n=\ell^2}. \qquad (14.7) \end{aligned}$$

The function

$$G(x,y,z) = \begin{cases} 0 & \text{if } x \text{ or } z \text{ is odd,} \\ (-1)^{(x+z)/2} F(x,y,z) & \text{if } x \text{ and } z \text{ are even} \end{cases}$$

is also odd in the variable x and even in the pair of variables y, z. Applying identity (14.6) to the function $G(x,y,z)$, we obtain

$$4 \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 0 \pmod 2}} (-1)^d F(\delta-2u, u+d, 2u+2d-\delta)$$

$$
\begin{aligned}
&- 2 \sum_{\substack{u^2+d\delta=n \\ d\equiv\delta \pmod 2}} (-1)^d F(d+\delta, u, d-\delta) \\
&= \left\{ 4 \sum_{\substack{1\le j\le 2\ell-1 \\ j\equiv 0 \pmod 2}} F(j,\ell,j) - 2 \sum_{j=-\ell+1}^{\ell-1} (-1)^{\ell+j} F(2\ell, j, 2j) \right\}_{n=\ell^2} \qquad (14.8)
\end{aligned}
$$

Subtracting (14.7) from (14.8) and dividing by 2, we obtain the important identity

$$
\begin{aligned}
&2 \sum_{u^2+d\delta=n} \varepsilon(d,\delta) \left(F(\delta-2u, u+d, 2u+2d-\delta) - \frac{1}{2} F(d+\delta, u, d-\delta) \right) \\
&= \left\{ 2 \sum_{j=1}^{2\ell-1} (-1)^{j-1} F(j,\ell,j) - \sum_{j=-\ell+1}^{\ell-1} (-1)^{\ell+j} F(2\ell, j, 2j) \right\}_{n=\ell^2} \qquad (14.9)
\end{aligned}
$$

where

$$
\varepsilon(d,\delta) = \begin{cases} 1 & \text{if } d \text{ and } \delta \text{ are even,} \\ -1 & \text{if } d \text{ or } \delta \text{ is odd.} \end{cases}
$$

The formula for $R_4(n)$ follows immediately from applying this identity to the function

$$
F(x,y,z) = xy^2.
$$

We obtain on the left side

$$
\begin{aligned}
&2 \sum_{u^2+d\delta=n} \varepsilon(d,\delta) \left((\delta-2u)(u+d)^2 - \frac{1}{2}(d+\delta)u^2 \right) \\
&= 2 \sum_{u^2+d\delta=n} \varepsilon(d,\delta) \left(d^2\delta + 2d\delta u + \frac{1}{2}\delta u^2 - 2u^3 - \frac{9}{2} du^2 - 2d^2 u \right) \\
&= 2 \sum_{u^2+d\delta=n} \varepsilon(d,\delta) \left(d(n-u^2) + \frac{1}{2}\delta u^2 - \frac{9}{2} du^2 \right) \qquad \text{(by (14.4))} \\
&= 2 \sum_{u^2+d\delta=n} \varepsilon(d,\delta) d(n-5u^2) - \sum_{u^2+d\delta=n} \varepsilon(d,\delta)(d-\delta)u^2 \\
&= \sum_{u^2<n} (n-5u^2) 2 \sum_{n-u^2=d\delta} \varepsilon(d,\delta) d \qquad \text{(by (14.5)).}
\end{aligned}
$$

If $n = \ell^2$, the right side of (14.9) is

$$
\begin{aligned}
2\ell^2 \sum_{j=1}^{2\ell-1} (-1)^{j-1} j - 2\ell \sum_{j=-\ell+1}^{\ell-1} (-1)^{\ell+j} j^2 &= 2\ell^3 - 4\ell \sum_{j=1}^{\ell-1} (-1)^{\ell-1-j} j^2 \\
&= 2\ell^3 - \frac{4\ell^2(\ell-1)}{2} \\
&= 2n,
\end{aligned}
$$

and so

$$\sum_{u^2<n}(n-5u^2)8\sum_{n-u^2=d\delta}\varepsilon(d,\delta)d=\{8n\}_{n=\ell^2}.$$

Define $\Phi(0)=1$ and

$$\Phi(n)=8\sum_{n=d\delta}\varepsilon(d,\delta)d$$

for $n\geq 1$. If n is not a square, then

$$\sum_{u^2\leq n}(n-5u^2)\Phi(n)=\sum_{u^2<n}(n-5u^2)\Phi(n)=0.$$

If n is a square and $n=\ell^2$, then

$$\begin{aligned}
&\sum_{u^2\leq n}(n-5u^2)\Phi(n)\\
&\quad=\sum_{u^2<n}(n-5u^2)\Phi(n)+\sum_{u=\pm\ell}(n-5u^2)\Phi(n)\\
&\quad=\sum_{u^2<n}(n-5u^2)\Phi(n)-8n\\
&\quad=0.
\end{aligned}$$

Therefore,

$$R_4(n)=8\sum_{n=d\delta}\varepsilon(d,\delta)d$$

for all positive integers n.

If n is odd and $n=d\delta$, then $\varepsilon(d,\delta)=1$ and

$$R_4(n)=8\sum_{d|n}d.$$

If n is even, then $n=2^a m$, where $a\geq 1$ and m is odd. Every divisor of n can be written uniquely in the form $2^b d$, where $0\leq b\leq a$ and $m=d\delta$. Then

$$\begin{aligned}
R_4(n) &= 8\sum_{m=d\delta}\sum_{b=0}^{a}\varepsilon(2^b d,2^{a-b}\delta)2^b d\\
&= 8\sum_{m=d\delta}\varepsilon(d,2^a\delta)d+8\sum_{m=d\delta}\varepsilon(2^a d,\delta)2^a d\\
&\quad+8\sum_{m=d\delta}\sum_{b=1}^{a-1}\varepsilon(2^b d,2^{a-b}\delta)2^b d\\
&= 8\sum_{m=d\delta}d+8\sum_{m=d\delta}2^a d-8\sum_{m=d\delta}\sum_{b=1}^{a-1}2^b d
\end{aligned}$$

$$= 8 \sum_{m=d\delta} d + 8 \sum_{m=d\delta} 2^a d - 8(2^a - 2) \sum_{m=d\delta} d$$
$$= 24 \sum_{m=d\delta} d$$
$$= 24 \sum_{\substack{d \mid n \\ d \equiv 1 \pmod 2}} d.$$

This completes the proof.

Exercises

1. Prove that $R_4(2^k) = 24$ for all $k \geq 1$. Find all representations of 2^k as a sum of four squares.

2. Prove that
$$\liminf_{n \to \infty} \frac{R_4(n)}{n^{\varepsilon}} = 0$$
for all $\varepsilon > 0$.

3. Compute $R_4(p^k)$ for all odd primes p and $k \geq 1$.

4. Prove that
$$\limsup_{n \to \infty} \frac{R_4(n)}{n} \geq 8.$$

5. Prove that
$$R_4(n) < 24 n \log n$$
for $n \geq 2$.

6. Prove that for every positive integer ℓ,
$$\sum_{j=1}^{\ell} (-1)^{\ell - j} j = \left[\frac{\ell + 1}{2} \right],$$
and so
$$\sum_{j=1}^{2\ell - 1} (-1)^{j-1} j = \ell.$$

7. Prove that for every positive integer ℓ,
$$\sum_{j=1}^{\ell} (-1)^{\ell - j} j^2 = \frac{\ell(\ell + 1)}{2},$$
and so
$$\sum_{j=1}^{2\ell - 1} (-1)^j (\ell - j)^2 = \ell - \ell^2.$$

14.5 Sums of Six Squares

In this section we obtain an explicit formula for $R_6(n)$. The idea is to apply identity (14.3) to the monomials x^3y and xy^3, and to manipulate the results so that we can find a function $\Phi(n)$ that satisfies the recursion formula

$$\sum_{|x|\leq\sqrt{n}} \left(n-7x^2\right)\Phi(n-x^2)=0.$$

Theorem 14.5 *Let n be a positive integer,*

$$n=2^a m,$$

where $a\geq 0$ and m is odd. Then

$$R_6(n)=4\left(4^{a+1}-(-1)^{(m-1)/2}\right)\sum_{m=d\delta}(-1)^{(\delta-1)/2}d^2.$$

As an example, we shall describe the representations of 5 as a sum of six squares. There are $2^5\binom{6}{5}=192$ representations as a sum of five terms $(\pm1)^2$. There are $2^2\binom{6}{1}\binom{5}{1}=120$ representations as a sum of $(\pm1)^2$ and $(\pm2)^2$. Thus, there are 312 representations of 5 as a sum of six squares.

We can also compute this number by applying Theorem 14.5 with $a=0$ and $m=5$. Then

$$R_6(5)=4\left(4^1-(-1)^{(5-1)/2}\right)(5^2+1)=4\cdot3\cdot26=312.$$

Proof. The function $f(x,y)=x^3y$ is odd in each of the variables x and y, and so we can apply (14.3) with $k_1=3$ and $k_2=1$. The left side of this identity is

$$\begin{aligned}
&\sum_{\substack{u^2+d\delta=n\\ \delta\equiv1 \pmod 2}}(-1)^{(\delta-1)/2}(\delta-2u)^3(u+d)\\
&=\sum_{\substack{u^2+d\delta=n\\ \delta\equiv1 \pmod 2}}(-1)^{(\delta-1)/2}(u\delta^3-6u^2\delta^2+12u^3\delta-8u^4+d\delta^3-6ud\delta^2\\
&\quad+12u^2d\delta-8u^3d)\\
&=\sum_{\substack{u^2+d\delta=n\\ \delta\equiv1 \pmod 2}}(-1)^{(\delta-1)/2}(d\delta^3-6u^2\delta^2+12u^2d\delta-8u^4)\\
&=\sum_{\substack{u^2+d\delta=n\\ \delta\equiv1 \pmod 2}}(-1)^{(\delta-1)/2}(\delta^2(n-7u^2)+4u^2(3n-5u^2)).
\end{aligned}$$

If $n=\ell^2$, then (by Exercise 3) the right side of the identity is

$$T_0(\ell)\;=\;(-1)^{\ell-1}\ell\sum_{k=1}^{\ell}(-1)^{k-1}(2k-1)^3$$

$$\begin{aligned} &= (-1)^{\ell-1}\ell(-1)^{\ell-1}(4\ell^3 - 3\ell) \\ &= 4\ell^4 - 3\ell^2 \\ &= 4n^2 - 3n. \end{aligned}$$

Therefore,

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta^2(n-7u^2) + 4u^2(3n-5u^2)) = \{4n^2 - 3n\}_{n=\ell^2}. \tag{14.10}$$

Next we apply (14.3) to the function $f(x,y) = xy^3$. The left side of the identity is

$$\begin{aligned} &\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta - 2u)(u+d)^3 \\ &= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(u^3\delta + 3u^2d\delta + 3ud^2\delta + d^3\delta - 2u^4 - 6u^3d \\ &\qquad - 6u^2d^2 - 2ud^3) \\ &= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d^3\delta - 6u^2d^2 + 3u^2d\delta - 2u^4) \\ &= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d^2(n-7u^2) + u^2(3n-5u^2)). \end{aligned}$$

If $n = \ell^2$, then (by Exercise 1) the right side of the identity is

$$\begin{aligned} T_0(\ell) &= (-1)^{\ell-1}\ell^3 \sum_{k=1}^{\ell} (-1)^{k-1}(2k-1) \\ &= \ell^4 \\ &= n^2. \end{aligned}$$

Multiplying by 4, we obtain

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(4d^2(n-7u^2) + 4u^2(3n-5u^2)) = \{4n^2\}_{n=\ell^2}. \tag{14.11}$$

Subtracting equation (14.10) from equation (14.11), we obtain

$$\begin{aligned} &\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(n-7u^2)(4d^2 - \delta^2) \\ &= \sum_{|u|<n} (n-7u^2) \sum_{\substack{d\delta=n-u^2 \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(4d^2-\delta^2) \\ &= \{3n\}_{n=\ell^2}. \end{aligned}$$

Let $\Phi(0) = 1$. For every positive integer n, define

$$\Phi(n) = 4 \sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(4d^2 - \delta^2).$$

If n is not a square, then

$$\sum_{|u|\leq n} (n - 7u^2)\Phi(n - u^2) = \sum_{|u|<n} (n - 7u^2)\Phi(n - u^2) = 0.$$

If $n = \ell^2$ is a square, then

$$\begin{aligned}
&\sum_{|u|\leq n} (n - 7u^2)\Phi(n - u^2) \\
&= \sum_{|u|<n} (n - 7u^2)\Phi(n - u^2) + (n - 7\ell^2)\Phi(0) + (n - 7(-\ell)^2)\Phi(0) \\
&= 12n - 12n \\
&= 0.
\end{aligned}$$

Therefore,

$$R_6(n) = \Phi(n) = 4 \sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(4d^2 - \delta^2).$$

We rewrite this equation as follows. Let $n = 2^a m$, where $a \geq 0$ and m is odd. Then δ is an odd divisor of n if and only if there exists a divisor d_1 of m such that $d = 2^a d_1$ and $m = d_1\delta$. Therefore,

$$\begin{aligned}
\sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} 4d^2 &= 4 \sum_{d_1\delta=m} (-1)^{(\delta-1)/2}(2^a d_1)^2 \\
&= 4^{a+1} \sum_{d_1\delta=m} (-1)^{(\delta-1)/2} d_1^2.
\end{aligned}$$

By Exercise 4, if m is odd and $d_1\delta = m$, then

$$(-1)^{(d-1)/2}(-1)^{(\delta-1)/2} = (-1)^{(m-1)/2}.$$

It follows that

$$\begin{aligned}
\sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}\delta^2 &= \sum_{d_1\delta=m} (-1)^{(\delta-1)/2}\delta^2 \\
&= \sum_{d\delta=m} (-1)^{(d-1)/2} d^2 \\
&= (-1)^{(m-1)/2} \sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2.
\end{aligned}$$

Therefore,

$$R_6(n) = \Phi(n) = 4\left(4^{a+1} - (-1)^{(m-1)/2}\right) \sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2.$$

This completes the proof.

Theorem 14.6 *For all positive integers n,*

$$\frac{3n^2}{2} < R_6(n) < 40n^2.$$

Proof. Let $n = 2^a m$, where $a \geq 0$ and m is odd. The infinite series $\zeta(2) = \sum_{k=1}^{\infty} k^{-2}$ converges, and $\zeta(2) < 2$ by Exercise 5. Then

$$\begin{aligned}
\sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2 &= m^2 \sum_{d\delta=m} \frac{(-1)^{(\delta-1)/2}}{\delta^2} \\
&\leq m^2 \sum_{d\delta=m} \frac{1}{\delta^2} \\
&< m^2 \sum_{k=1}^{\infty} \frac{1}{k^2} \\
&< 2m^2
\end{aligned}$$

and

$$\begin{aligned}
4^{a+1} - (-1)^{(m-1)/2} &\leq 4 \cdot 4^a + 1 \\
&\leq 5\left(2^a\right)^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
R_6(n) &= 4\left(4^{a+1} - (-1)^{(m-1)/2}\right) \sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2 \\
&\leq 4 \cdot 5\left(2^a\right)^2 2m^2 \\
&= 40n^2.
\end{aligned}$$

This gives the upper bound.

To obtain a lower bound, we have

$$\begin{aligned}
\sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2 &= m^2 \sum_{d\delta=m} \frac{(-1)^{(\delta-1)/2}}{\delta^2} \\
&\geq m^2 \left(1 - \sum_{\substack{\delta \mid m \\ \delta > 1}} \frac{1}{\delta^2}\right) \\
&> m^2 \left(1 - \sum_{k=1}^{\infty} \frac{1}{(2k+1)^2}\right) \\
&> \frac{m^2}{2}
\end{aligned}$$

by Exercise 6. Also,

$$\begin{aligned} 4^{a+1} - (-1)^{(m-1)/2} &\geq 4 \cdot 4^a - 1 \\ &\geq 3\left(2^a\right)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} R_6(n) &= 4\left(4^{a+1} - (-1)^{(m-1)/2}\right) \sum_{d\delta=m} (-1)^{(\delta-1)/2} d^2 \\ &\geq 3\left(2^a\right)^2 \frac{m^2}{2} \\ &= \frac{3n^2}{2}. \end{aligned}$$

This completes the proof.

Exercises

1. Find all representations of 6 as a sum of 6 squares.
2. Find all representations of 10 as a sum of 6 squares.
3. Prove that for every positive integer m,
$$\sum_{j=1}^{\ell} (-1)^{\ell-j}(2j-1)^3 = (4\ell^3 - 3\ell).$$
4. Prove that if m is odd and $d\delta = m$, then
$$(-1)^{(d-1)/2}(-1)^{(\delta-1)/2} = (-1)^{(m-1)/2}.$$
5. Prove that
$$\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2} < 2.$$
Hint: $k^{-2} < \int_{k-1}^{k} x^{-2} dx$ for $k \geq 2$.
6. Prove that
$$\sum_{k=1}^{\infty} \frac{1}{(2k+1)^2} < \frac{1}{2}.$$
Hint: $4(2k+1)^{-2} < k^{-2}$.
7. Use the fact that $\zeta(2) = \pi^2/6$ to prove that
$$\sum_{k=1}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{24} - 1 = 0.23\ldots.$$

14.6 Sums of Eight Squares

Theorem 14.7 *Let n be a positive integer. If n is odd, then*

$$R_8(n) = 16 \sum_{d|n} d^3.$$

If n is even and $n = 2^a m$, where $a \geq 1$ and m is odd, then

$$R_8(n) = \frac{16(8^{a+1} - 15)}{7} \sum_{d|m} d^3.$$

Proof. We shall apply Liouville's identity (Theorem 13.1) to the three polynomials $(-1)^y xy^4, (-1)^y xy^3(2y - z)$, and $(-1)^y xy^2$.

Inserting $(-1)^y xy^4$ into Liouville's identity, we find that the first term on the left is

$$\begin{aligned} & 2 \sum_{u^2+d\delta=n} (-1)^{u+d}(\delta - 2u)(u + d)^4 \\ = \;& 2 \sum_{u^2+d\delta=n} (-1)^{u+d} \left(d^4\delta - 8u^2d^3 + u^4\delta - 8u^4 d + 6u^2d^2\delta\right) \\ = \;& 2 \sum_{u^2+d\delta=n} (-1)^{u+d} \left(d^3(n - 9u^2) + u^4(\delta - 14d) + 6nu^2 d\right). \end{aligned}$$

The second term on the left side of the identity is

$$\sum_{u^2+d\delta=n} (-1)^u (d + \delta)u^4 = 2 \sum_{u^2+d\delta=n} (-1)^u du^4.$$

If $n = \ell^2$, then

$$2T_1(\ell) = (-1)^\ell 2\ell^4 \sum_{j=1}^{2\ell-1} j = (-1)^\ell (4\ell^6 - 2\ell^5)$$

by Exercise 2, and

$$\begin{aligned} T_2(\ell) &= 2\ell \sum_{j=-\ell+1}^{\ell-1} (-1)^j j^4 = 4\ell \sum_{j=1}^{\ell-1} (-1)^j j^4 \\ &= (-1)^{\ell-1} \left(2\ell^5 - 4\ell^4 + 2\ell^2\right), \end{aligned}$$

and so the right side of Liouville's identity is

$$2T_1(\ell) - T_2(\ell) = (-1)^\ell (4\ell^6 - 4\ell^4 + 2\ell^2) = (-1)^n (4n^3 - 4n^2 + 2n).$$

Dividing by 2, we obtain

$$\sum_{u^2+d\delta=n} (-1)^{u+d} d^3(n-9u^2) + \sum_{u^2+d\delta=n} (-1)^u u^4 \left((-1)^d(\delta-14d)-d\right)$$
$$+ 6n \sum_{u^2+d\delta=n} (-1)^{u+d} du^2 = \{(-1)^n(2n^3-2n^2+n)\}_{n=\ell^2} \quad (14.12)$$

Next we consider the polynomial $(-1)^y xy^3(2y-z)$. The first term on the left side of Liouville's formula is

$$\begin{aligned}
&2 \sum_{(u,d,\delta)\in\mathcal{S}(n)} (-1)^{u+d}(\delta-2u)(u+d)^3\delta \\
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}(n)} (-1)^{u+d} \left(3d\delta^2u^2 + d^3\delta^2 - 2\delta u^4 - 6d^2\delta u^2\right) \\
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}(n)} (-1)^{u+d}(3\delta u^2(n-u^2) + d(n-u^2)^2 \\
&\quad - 2\delta u^4 - 6du^2(n-u^2)) \\
&= 2 \sum_{(u,d,\delta)\in\mathcal{S}(n)} (-1)^{u+d} \left(nu^2(3\delta-8d) + u^4(7d-5\delta) + n^2d\right).
\end{aligned}$$

The second term on the left is

$$\begin{aligned}
\sum_{u^2+d\delta=n} (-1)^u(d+\delta)u^3(2u-d+\delta) &= 2 \sum_{u^2+d\delta=n} (-1)^u(d+\delta)u^4 \\
&= 4 \sum_{u^2+d\delta=n} (-1)^u du^4.
\end{aligned}$$

If $n=\ell^2$, then

$$\begin{aligned}
2T_1(\ell) &= 2\sum_{j=1}^{2\ell-1}(-1)^\ell j\ell^3(2\ell-j) \\
&= (-1)^\ell 4\ell^4 \sum_{j=1}^{2\ell-1} j - (-1)^\ell 2\ell^3 \sum_{j=1}^{2\ell-1} j^2 \\
&= \frac{(-1)^n 2(4n^3-n^2)}{3}
\end{aligned}$$

and

$$T_2(\ell) = 0.$$

Therefore,

$$
\begin{aligned}
2 \sum_{u^2+d\delta=n} & (-1)^{u+d} \left(nu^2(3\delta - 8d) + u^4(7d - 5\delta) + n^2 d\right) \\
& - 4 \sum_{u^2+d\delta=n} (-1)^u du^4 \\
= & \left\{ \frac{(-1)^n 2(4n^3 - n^2)}{3} \right\}_{n=\ell^2},
\end{aligned}
$$

or, equivalently,

$$
\begin{aligned}
3 \sum_{u^2+d\delta=n} & (-1)^u u^4 \left((-1)^d (7d - 5\delta) - 2d\right) \\
& + 3n \sum_{u^2+d\delta=n} (-1)^{u+d} u^2 (3\delta - 8d) + 3n^2 \sum_{u^2+d\delta=n} (-1)^{u+d} d \\
= & \left\{(-1)^n (4n^3 - n^2)\right\}_{n=\ell^2}. \qquad (14.13)
\end{aligned}
$$

For every positive integer n we have

$$
\begin{aligned}
\sum_{u^2+d\delta=n} & (-1)^u u^4 \left((-1)^d (\delta - 14d) - d\right) \\
& + 3 \sum_{u^2+d\delta=n} (-1)^u u^4 \left((-1)^d (7d - 5\delta) - 2d\right) \\
= & \; 7 \sum_{u^2<n} (-1)^u u^4 \sum_{n-u^2=d\delta} \left((-1)^d (d - 2\delta) - d\right) \\
= & \; 0
\end{aligned}
$$

by Exercise 3. Adding equations (14.12) and (14.13), we obtain

$$
\begin{aligned}
\sum_{u^2+d\delta=n} & (-1)^{u+d} d^3 (n - 9u^2) + 9n \sum_{u^2+d\delta=n} (-1)^{u+d} u^2 (\delta - 2d) \\
& + 3n^2 \sum_{u^2+d\delta=n} (-1)^{u+d} d = \{(-1)^n (6n^3 - 3n^2 + n)\}_{n=\ell^2}. \quad (14.14)
\end{aligned}
$$

Finally, we consider the polynomial $(-1)^y xy^2$. The left side of Liouville's identity is

$$
\begin{aligned}
2 \sum_{u^2+d\delta=n} & (-1)^{u+d} (\delta - 2u)(u + d)^2 - \sum_{u^2+d\delta=n} (-1)^u (d + \delta) u^2 \\
= & \; 2 \sum_{u^2+d\delta=n} (-1)^{u+d} (u^2 (\delta - 5d) + nd) - 2 \sum_{u^2+d\delta=n} (-1)^u du^2.
\end{aligned}
$$

If $n = \ell^2$, then

$$
2T_1(\ell) - T_2(\ell) = (-1)^\ell (4\ell^4 - 2\ell^2) = (-1)^n (4n^2 - 2n).
$$

Multiplying by $3n/2$, we obtain

$$\begin{aligned} 3n & \sum_{u^2+d\delta=n} (-1)^u u^2 \left((-1)^d(\delta-5d)-d\right) + 3n^2 \sum_{u^2+d\delta=n} (-1)^{u+d} d \\ &= 9n \sum_{u^2+d\delta=n} (-1)^{u+d} u^2(\delta-2d) + 3n^2 \sum_{u^2+d\delta=n} (-1)^{u+d} d \\ &= \{(-1)^n(6n^3-3n^2)\}_{n=\ell^2}, \end{aligned} \tag{14.15}$$

since

$$\sum_{n-u^2=d\delta} \left((-1)^d(\delta-5d)-d\right) = 3 \sum_{n-u^2=d\delta} (-1)^d(\delta-2d)$$

by Exercise 3. Subtracting (14.15) from (14.14), we obtain

$$\sum_{u^2+d\delta=n} (-1)^{u+d} d^3(n-9u^2) = \{(-1)^n n\}_{n=\ell^2}.$$

We define the function $\Phi(n)$ as follows:

$$\Phi(0) = 1$$

and

$$\Phi(n) = 16(-1)^n \sum_{d|n} (-1)^d d^3$$

for every positive integer n. If n is not a square, then

$$\begin{aligned} \sum_{u^2 \le n} (n-9u^2)\Phi(n-u^2) &= \sum_{u^2<n} (n-9u^2)\Phi(n-u^2) \\ &= 16 \sum_{u^2<n} (n-9u^2)(-1)^{n-u^2} \sum_{n-u^2=d\delta} (-1)^d d^3 \\ &= 16(-1)^n \sum_{u^2<n} (n-9u^2)(-1)^u \sum_{n=d\delta} (-1)^d d^3 \\ &= 0. \end{aligned}$$

If $n = \ell^2$, then

$$\begin{aligned} &\sum_{u^2 \le n} (n-9u^2)\Phi(n-u^2) \\ &\quad = \sum_{u^2<n} (n-9u^2)\Phi(n-u^2) + \sum_{u=\pm\ell} (n-9u^2)\Phi(n-u^2) \\ &\quad = 16(-1)^n \sum_{u^2<n} (n-9u^2)(-1)^u \sum_{n-u^2=d\delta} (-1)^d d^3 - 16n \\ &\quad = 16(-1)^n \{(-1)^n n\}_{n=\ell^2} - 16n \\ &\quad = 0. \end{aligned}$$

The recursion formula (14.2) implies that

$$R_8(n) = \Phi(n).$$

We can rewrite the expression for $R_8(n)$ as follows. Let $n = 2^a m$, where $a \geq 0$ and m is odd. The odd divisors of n are precisely the divisors of m. The even divisors of n are the numbers of the form $2^b d$, where $1 \leq b \leq a$ and d is a divisor of m. Then

$$\begin{aligned}
\Phi(n) &= 16(-1)^n \sum_{n=d\delta} (-1)^d d^3 \\
&= 16(-1)^n \left(\sum_{b=1}^{a} \sum_{d|m} (2^b d)^3 - \sum_{d|m} d^3 \right) \\
&= 16(-1)^n \left(\sum_{b=1}^{a} 8^b - 1 \right) \sum_{d|m} d^3 \\
&= 16(-1)^n \left(\frac{8^{a+1} - 15}{7} \right) \sum_{d|m} d^3.
\end{aligned}$$

This completes the proof.

Exercises

1. Prove that for every positive integer n,

$$16n^3 < R_8(n) < \left(\frac{128\zeta(3)}{7} \right) n^3,$$

   where $\zeta(3) = \sum_{k=1}^{\infty} k^{-3}$.

2. Prove that for every positive integer ℓ,

$$\sum_{j=1}^{\ell-1} (-1)^j j^4 = (-1)^{\ell-1} \left(\frac{\ell^4 - 2\ell^3 + \ell}{2} \right).$$

3. Prove that

$$\sum_{n=d\delta} \left((-1)^d (d - 2\delta) - d \right) = 0$$

   for every positive integer n.

14.7 Sums of Ten Squares

We shall determine the number of representations of an integer as a sum of ten squares. In this case the formula for $R_{10}(n)$ contains two terms. The

first is a divisor function, that is, a sum over divisors of n, and the second is a sum over representations of n as a sum of two squares.

Theorem 14.8 *Let n be a positive integer,*

$$n = 2^a m,$$

where $a \geq 0$ and m is odd. Then

$$\begin{aligned} R_{10}(n) &= \frac{4}{5}\left(16^{a+1} + (-1)^{(m-1)/2}\right) \sum_{m=d\delta} (-1)^{(\delta-1)/2} d^4 \\ &\quad + \frac{16}{5} \sum_{n=v^2+w^2} \left(v^4 - 3v^2w^2\right). \end{aligned}$$

As an example, we list the representations of 5 as a sum of ten squares. There are $2^5\binom{10}{5} = 32 \cdot 252 = 8064$ representations as a sum of five terms of the form $(\pm 1)^2$. There are $2^2\binom{10}{1}\binom{9}{1} = 360$ representations as a sum of the integers $(\pm 1)^2$ and $(\pm 2)^2$. Thus, there are 8424 representations. By Theorem 14.8, with $n = m = 5$ and $a = 0$, we have

$$\begin{aligned} R_{10}(5) &= \frac{4}{5}(16+1)(5^4+1) + \frac{16}{5} \sum_{5=x^2+y^2} \left(x^4 - 3x^2y^2\right) \\ &= \frac{42568}{5} + \frac{16}{5}(4(2^4 - 3\cdot 2^2) + 4(1^4 - 3 \cdot 2^2)) \\ &= \frac{42568}{5} - \frac{448}{5} \\ &= 8424. \end{aligned}$$

Proof. By Theorem 14.2, it suffices to find a function $\Phi(n)$ such that $\Phi(0) = 1$ and

$$\sum_{|x| \leq \sqrt{n}} \left(n - 11x^2\right) \Phi(n - x^2) = 0$$

for every positive integer n.

We begin by applying identity (14.3) to each of the monomials x^5y, x^3y^3, and xy^5. With $f(x, y) = x^5y$, we obtain

$$\begin{aligned} &\sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta - 2u)^5(u+d) \\ &= \sum_{\substack{u^2+d\delta=n \\ \delta \equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} \\ &\quad \times \left(\sum_{\substack{0 \leq k \leq 5 \\ k \equiv 1 \pmod 2}} \binom{5}{k} (-2)^k \delta^{5-k} u^{k+1} + \sum_{\substack{0 \leq k \leq 5 \\ k \equiv 0 \pmod 2}} \binom{5}{k} (-2)^k d\delta^{5-k} u^k \right) \end{aligned}$$

$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d\delta^5 - 10\delta^4u^2 + 40d\delta^3u^2 - 80\delta^2u^4$$
$$+ 80d\delta u^4 - 32u^6)$$
$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta^4(n-u^2) - 10\delta^4u^2 + 40\delta^2u^2(n-u^2)$$
$$- 80\delta^2u^4 + 16u^4(5n-5u^2) - 32u^6)$$
$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta^4(n-11u^2) + 40\delta^2u^2(n-3u^2)$$
$$+ 16u^4(5n-7u^2))$$
$$= \left\{\ell\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)^5\right\}_{n=\ell^2}$$
$$= \left\{16n^3 - 40n^2 + 25n\right\}_{n=\ell^2}$$

by Exercise 4.

Applying (14.3) with $f(x,y) = x^3y^3$, we obtain

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta-2u)^3(u+d)^3$$
$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(3d\delta^3u^2 + 12d^3\delta u^2 - 6\delta^2u^4 - 24d^2u^4 + d^3\delta^3$$
$$- 18d^2\delta^2u^2 + 36d\delta u^4 - 8u^6)$$
$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}((3\delta^2u^2 + 12d^2u^2)(n-u^2)$$
$$- (3\delta^2u^2 + 12d^2u^2)2u^2(d\delta - 2u^2)((d\delta-2u^2)^2 - 12d\delta u^2))$$
$$= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}((3\delta^2u^2 + 12d^2u^2)(n-3u^2)$$
$$+ (n-3u^2)^3 - 12u^2(n-u^2)(n-3u^2))$$
$$= \left\{\ell^3\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)^3\right\}_{n=\ell^2}$$
$$= \left\{4n^3 - 3n^2\right\}_{n=\ell^2}$$

by Exercise 3 in Section 14.5.

Applying (14.3) with $f(x,y) = xy^5$, we obtain

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta-2u)(u+d)^5$$

$$\begin{aligned}
&= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d^5\delta - 10d^4u^2 + 10d^3\delta u^2 - 20d^2u^4 \\
&\quad + 5d\delta u^4 - 2u^6) \\
&= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d^4(n-11u^2) + 10d^2u^2(n-3u^2) \\
&\quad + u^4(5n-7u^2)) \\
&= \left\{\ell^5\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)\right\}_{n=\ell^2} \\
&= \{n^3\}_{n=\ell^2}
\end{aligned}$$

by Exercise 1 in Section 14.3.

The upshot of this analysis is the following three identities:

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(\delta^4(n-11u^2) + 40\delta^2u^2(n-3u^2) + 16u^4(5n-7u^2)) = \left\{16n^3 - 40n^2 + 25n\right\}_{n=\ell^2}, \tag{14.16}$$

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}((3\delta^2u^2 + 12d^2u^2)(n-3u^2) + (n-3u^2)^3 - 12u^2(n-u^2)(n-3u^2)) = \left\{4n^3 - 3n^2\right\}_{n=\ell^2}, \tag{14.17}$$

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(d^4(n-11u^2) + 10d^2u^2(n-3u^2) + u^4(5n-7u^2)) = \{n^3\}_{n=\ell^2}. \tag{14.18}$$

We shall eliminate the terms

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}d^2u^2(n-3u^2)$$

and

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}\delta^2u^2(n-3u^2)$$

from these equations as follows: Multiply equation (14.18) by 16 and add to equation (14.16), then multiply equation (14.17) by 40/3 and subtract. We obtain

$$\sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}(n-11u^2)(16d^4+\delta^4) + \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2}$$

$$\times\left(160u^2(n-u^2)(n-3u^2)+32u^4(5n-7u^2)\frac{40}{3}(n-3u^2)^3\right)$$
$$= \left\{25n-\frac{64n^3}{3}\right\}_{n=\ell^2}.$$

Let $P(n)$ denote the first sum in this equation, and let $Q(n)$ denote the second sum. Then

$$P(n)-\{25n\}_{n=\ell^2}+Q(n)+\left\{\frac{64n^3}{3}\right\}_{n=\ell^2}=0.$$

For positive integers n we define the function $\varphi(n)$ by

$$\varphi(n)=\sum_{\substack{n=d\delta\\ d,\delta\geq 1\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}(16d^4+\delta^4).$$

Let

$$\varphi(0)=\frac{5}{4}.$$

Then

$$\begin{aligned}
P(n) &= \sum_{\substack{n=u^2+d\delta\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}(n-11u^2)(16d^4+\delta^4)\\
&= \sum_{u^2<n}(n-11u^2)\sum_{\substack{n=d\delta\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}(16d^4+\delta^4)\\
&= \sum_{u^2<n}(n-11u^2)\varphi(n-u^2).
\end{aligned}$$

If $n=\ell^2$ is a square, then

$$\begin{aligned}
\sum_{u=\pm\ell}(n-11u^2)\varphi(n-u^2) &= (n-11\ell^2)\varphi(0)+(n-11(-\ell)^2)\varphi(0)\\
&= (-20n)\frac{5}{4}\\
&= -25n,
\end{aligned}$$

and so

$$P(n)-\{25n\}_{n=\ell^2}=\sum_{u^2\leq n}(n-11u^2)\varphi(n-u^2).$$

Recall the formula for the number of representations of an integer as the sum of two squares:

$$R_2(n)=4\sum_{\substack{\delta|n\\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}.$$

Then

$$\begin{aligned}
Q(n) &= \sum_{\substack{u^2+d\delta=n \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} \\
&\quad \times \left(160u^2(n-u^2)(n-3u^2)+32u^4(5n-7u^2)-\frac{40}{3}(n-3u^2)^3\right) \\
&= \sum_{u^2<n} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&\quad \times 4 \sum_{\substack{\delta\mid(n-u^2) \\ \delta\equiv 1 \pmod 2}} (-1)^{(\delta-1)/2} \\
&= \sum_{u^2<n} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&\quad \times R_2(n-u^2).
\end{aligned}$$

If $n=\ell^2$, then $R_2(n-\ell^2)=R_2(0)=1$ and

$$\begin{aligned}
&\sum_{u=\pm\ell} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&\qquad \times R_2(n-u^2) \\
&\quad = \frac{64n^3}{3},
\end{aligned}$$

and so

$$\begin{aligned}
&Q(n)+\left\{\frac{64n^3}{3}\right\}_{n=\ell^2} \\
&= \sum_{u^2\le n} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&\quad \times R_2(n-u^2) \\
&= \sum_{u^2\le n} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&\quad \times \sum_{n-u^2=v^2+w^2} 1 \\
&= \sum_{n=u^2+v^2+w^2} \left(40u^2(n-u^2)(n-3u^2)+8u^4(5n-7u^2)-\frac{10}{3}(n-3u^2)^3\right) \\
&= \sum_{n=u^2+v^2+w^2} (40u^2(v^2+w^2)(v^2+w^2-2u^2)+8u^4(5v^2+5w^2-2u^2) \\
&\quad -\frac{10}{3}(v^2+w^2-2u^2)^3)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{n=u^2+v^2+w^2} \left(\frac{32u^6}{3} - \frac{10v^6}{3} - \frac{10w^6}{3} \right) + \sum_{n=u^2+v^2+w^2} 120u^2v^2w^2 \\
&\quad + \sum_{n=u^2+v^2+w^2} \left(60u^4v^2 - 80u^2v^4 + 60u^4w^2 - 80u^2w^4 - 10v^4w^2 - 10v^2w^4 \right) \\
&= 4 \sum_{n=u^2+v^2+w^2} \left(u^6 - 15u^4v^2 + 30u^2v^2w^2 \right).
\end{aligned}$$

The simple form of the last equation arises from a symmetry argument: If $h(u,v,w)$ is any function and σ is any permutation of u, v, and w, then

$$\sum_{n=u^2+v^2+w^2} h(u,v,w) = \sum_{n=u^2+v^2+w^2} h(\sigma(u), \sigma(v), \sigma(w)).$$

For every nonnegative integer n we define the function

$$\psi(n) = \sum_{n=v^2+w^2} (v^4 - 3v^2w^2).$$

Then $\psi(0) = 0$, $\psi(1) = 2$, $\psi(2) = -8, \ldots$, and

$$\begin{aligned}
&\sum_{u^2 \le n} (n - 11u^2)\psi(n - u^2) \\
&\quad = \sum_{u^2 \le n} (n - 11u^2) \sum_{n-u^2=v^2+w^2} (v^4 - 3v^2w^2) \\
&\quad = \sum_{n=u^2+v^2+w^2} (n - 11u^2)(v^4 - 3v^2w^2) \\
&\quad = \sum_{n=u^2+v^2+w^2} (v^2 + w^2 - 10u^2)(v^4 - 3v^2w^2) \\
&\quad = \sum_{n=u^2+v^2+w^2} (v^6 - 2v^4w^2 - 3v^2w^4 - 10u^2v^4 + 30u^2v^2w^2) \\
&\quad = \sum_{n=u^2+v^2+w^2} (u^6 - 15u^4v^2 + 30u^2v^2w^2) \qquad \text{by (14.5).}
\end{aligned}$$

Therefore,

$$Q(n) + \left\{ \frac{64n^3}{3} \right\}_{n=\ell^2} = 4 \sum_{u^2 \le n} (n - 11u^2)\psi(n - u^2).$$

We define

$$\Phi(n) = \frac{4(\varphi(n) + 4\psi(n))}{5}.$$

Then

$$\Phi(0) = 1$$

and

$$\sum_{u^2\leq n}(n-11u^2)\Phi(n-u^2)=0$$

for all positive integers n. It follows that

$$\begin{aligned} R_{10}(n) &= \frac{4}{5}\left(\varphi(n)+4\psi(n)\right) \\ &= \frac{4}{5}\sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}(16d^4+\delta^4)+\frac{16}{5}\sum_{n=v^2+w^2}(v^4-3v^2w^2). \end{aligned}$$

Let $n=2^a m$, where m is odd and $a\geq 0$. Since $n=d\delta$ with δ odd if and only if d is of the form $d=2^a d_1$, where d_1 is a divisor of m, then it follows that

$$\sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}16d^4=16^{a+1}\sum_{d_1\delta=m}(-1)^{(\delta-1)/2}d_1^4.$$

Moreover, if $m=d_1\delta$, then

$$(-1)^{(m-1)/2}=(-1)^{(d_1-1)/2}(-1)^{(\delta-1)/2}$$

and

$$\begin{aligned} \sum_{\substack{d\delta=n \\ \delta\equiv 1 \pmod 2}}(-1)^{(\delta-1)/2}\delta^4 &= \sum_{d_1\delta=m}(-1)^{(\delta-1)/2}\delta^4 \\ &= \sum_{d_1\delta=m}(-1)^{(d_1-1)/2}d_1^4 \\ &= (-1)^{(m-1)/2}\sum_{d_1\delta=m}(-1)^{(\delta-1)/2}d_1^4. \end{aligned}$$

This completes the proof.

Exercises

1. Compute $R_{10}(n)$ for $n=1,\ldots,10$.

2. Find all representations of 10 as a sum of 10 squares.

3. Find all representations of 6 as a sum of 10 squares.

4. Prove that for every positive integer ℓ,

$$\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)^5=16\ell^5-40\ell^3+25\ell.$$

5. Evaluate the sum
$$\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)^2.$$

6. Evaluate the sum
$$\sum_{j=1}^{\ell}(-1)^{\ell-j}(2j-1)^4.$$

7. A *Gaussian integer* is a complex number $v+wi$, where v and w are ordinary integers. The *norm* of the Gaussian integer $v+wi$ is $N(v+wi)=v^2+w^2$. Prove that
$$\sum_{n=v^2+w^2}(v^4-3v^2w^2)=\frac{1}{2}\sum_{N(v+wi)=n}(v+wi)^4.$$

14.8 Notes

Liouville's identity, applied to "appropriate" polynomials and rearranged, gives formulae for the number of representations of an integer as the sum of an even number of squares. Our manipulations evolved the old-fashioned way, by hand with pencil and paper, but almost certainly it is possible today to do this more efficiently with human-assisted computer algebra systems. It would be a useful exercise to derive formulae for $R_s(n)$ for even numbers $s\geq 12$ using software such as Maple or Mathematica.

The proofs in this chapter are based on Venkov's exposition [149] of Liouville's method. Analytic proofs of these results can be found in the books of Grosswald [43], Knopp [81], and Rademacher [119]. An interesting discussion of the problem of sums of squares appears in Hardy's book *Ramanujan* [52, Chapter IX].

Iwaniec [74] considers the more general problem of the number of representations of an integer n by a positive definite quadratic form $Q(x_1,\ldots,x_s)$. We denote the representation number by $r_Q(n)$. This is the Fourier coefficient of the *theta function*
$$\theta_Q(z)=\sum_{(x_1,\ldots,x_s)\in\mathbf{Z}^s}e^{2\pi iQ(x_1,\ldots,x_s)z}=\sum_{n=0}^{\infty}r_Q(n)e^{2\pi inz},$$
and
$$\theta_Q(z)=E_Q(z)+F_Q(z),\tag{14.19}$$
where $E_Q(z)$ is an Eisenstein series and $F_Q(z)$ is a cusp form.

In this chapter we considered the positive definite quadratic form
$$Q(x_1,\ldots,x_s)=x_1^2+\cdots+x_s^2.$$

If s is even and $s \leq 8$, then the cusp form in (14.19) is zero and $r_s(n)$ is the coefficient of an Eisenstein series. If s is even and $s \geq 10$, then the cusp form in (14.19) is nonzero, and the main term in $r_s(n)$ is the coefficient of an Eisenstein series and the remainder term is the coefficient of a cusp form. In this case, Liouville's formulae might provide a method to compute the coefficients of cusp forms.

15
Partition Asymptotics

15.1 The Size of $p(n)$

A *partition* of n is a representation of n as a sum of positive integers. The order of the summands does not matter. We often write the partition in the form

$$n = a_1 + a_2 + \cdots + a_k,$$

where

$$a_1 \geq a_2 \geq \cdots \geq a_k \geq 1.$$

For example, the partitions of 5 are

$$\begin{aligned}
&5, \\
&4+1, \\
&3+2, \\
&3+1+1, \\
&2+2+1, \\
&2+1+1+1, \\
&1+1+1+1+1.
\end{aligned}$$

The *unrestricted partition function* $p(n)$ counts the number of partitions of the positive integer n. Thus, $p(5) = 7$. This function is strictly increasing, and satisfies the asymptotic formula

$$p(n) \sim \frac{e^{c_0\sqrt{n}}}{(4\sqrt{3})n}, \tag{15.1}$$

where

$$c_0 = \pi\sqrt{\frac{2}{3}} = 2\sqrt{\frac{\pi^2}{6}} = 2.565\ldots.$$

It follows that

$$\log p(n) \sim c_0\sqrt{n}. \tag{15.2}$$

Hardy and Ramanujan [58] and Uspensky [146] independently discovered this result; their proofs used complex variables and modular functions. Erdős later found an elementary proof of (15.1). The idea of Erdős's proof is simply to apply induction to the recursion formula (Theorem 15.1)

$$np(n) = \sum_{\substack{kv\leq n\\ k,v\geq 1}} vp(n-kv). \tag{15.3}$$

The proof, however, is difficult; it is "elementary" only in the technical sense that it does not require complex analysis. We shall use Erdős's method to obtain (15.2). The determination of the asymptotics of partition functions is our third problem in additive number theory.

Let A be a nonempty set of positive integers, and let $d = \gcd(A)$. For every positive integer n, the partition function $p_A(n)$ counts the number of partitions of n into parts belonging to A. We define $p_A(0) = 1$ for all sets A. We would like to understand the asymptotic behavior of $p_A(n)$. For example, if A is the set of odd positive integers, then $p_A(n)$ is the number of partitions of n into odd parts, and $\log p_A(n) \sim \pi\sqrt{n/3}$.

If $d = \gcd(A) > 1$, we consider the set $A' = \{a/d : a \in A\}$. Then $\gcd(A') = 1$, and

$$p_A(n) = \begin{cases} 0 & \text{if } n \not\equiv 0 \pmod{d},\\ p_{A'}(n/d) & \text{if } n \equiv 0 \pmod{d}.\end{cases}$$

Thus, it suffices to consider only partition functions for sets A such that $\gcd(A) = 1$.

We do this in two significant cases. In the first, A is a finite set of integers with $|A| = k$ and $\gcd(A) = 1$. We shall prove that

$$p_A(n) = \left(\frac{1}{\prod_{a\in A} a}\right)\frac{n^{k-1}}{(k-1)!} + O\left(n^{k-2}\right).$$

In the second, A is a set of integers of positive density $d(A) = \alpha$ with $\gcd(A) = 1$. We shall prove that

$$\log p_A(n) \sim c_0\sqrt{\alpha n}. \tag{15.4}$$

We shall also prove an inverse theorem: If A is a set of positive integers whose partition function satisfies (15.4) for some $\alpha > 0$, then $\gcd(A) = 1$ and A has density α.

We begin by proving the recursion formula (15.3).

Theorem 15.1 *For every positive integer n,*

$$np(n) = \sum_{\substack{kv \leq n \\ k,v \geq 1}} vp(n - kv).$$

Proof. The parts in a partition of n are positive integers v not exceeding n. The number of partitions of n with at least one part equal to v is $p(n-v)$. For any positive integer k, the number of partitions of n with at least k parts equal to v is $p(n - kv)$, and so the number of partitions of n with exactly k parts equal to v is $p(n - kv) - p(n - (k + 1)v)$. Therefore, the number of parts equal to v that occur in all partitions of n is

$$\sum_{k \geq 1} k(p(n - kv) - p(n - (k + 1)v)) = \sum_{k \geq 1} p(n - kv).$$

We list the $p(n)$ partitions of n as follows:

$$\begin{aligned}
n &= a_{1,1} + a_{1,2} + \cdots + a_{1,k_1}, \\
n &= a_{2,1} + a_{2,2} + \cdots + a_{2,k_2}, \\
n &= a_{3,1} + a_{3,2} + \cdots + a_{3,k_3}, \\
&\vdots \\
n &= a_{p(n),1} + a_{p(n),2} + \cdots + a_{p(n),k_{p(n)}}.
\end{aligned}$$

Adding the $p(n)$ rows of this array, we obtain

$$\begin{aligned}
np(n) &= \sum_{i=1}^{p(n)} \sum_{j=1}^{k_i} a_{i,j} \\
&= \sum_{v=1}^{n} v \sum_{a_{i,j}=v} 1 \\
&= \sum_{v=1}^{n} v \sum_{k \geq 1} p(n - kv) \\
&= \sum_{\substack{kv \leq n \\ k,v \geq 1}} vp(n - kv).
\end{aligned}$$

This completes the proof. □

Exercises

1. Compute $p(n)$ for $n = 1, 2, 3, 4$.

2. Let $q(n)$ denote the number of partitions of n into distinct parts. Let A be the set of odd numbers and $p_A(n)$ the number of partitions of n into not necessarily distinct odd parts. Compute $p(6), q(6)$, and $p_A(6)$.

3. Compute $p(7), q(7)$, and $p_A(7)$.

4. Use the recursion formula (15.3) to compute $p(8)$.

5. Let $A = \{1\} \cup \{2n : n \geq 1\}$. Prove that
$$p_A(2n) = p_A(2n+1)$$
for all nonnegative integers n.

6. Prove that if $p_A(n) \geq 1$ and $p_A(n_0) \geq 1$, then $p_A(n) \leq p_A(n+n_0)$.

7. Let A be a nonempty set of positive integers, and let $a_1 \in A$. Prove that the partition function $p_A(n)$ is increasing in every congruence class modulo a_1, that is,
$$p_A(n) \leq p_A(n+a_1)$$
for every positive integer n.

   Prove that for every real number $x \geq a_1$ there exists an integer u such that
$$x - a_1 < u \leq x$$
and
$$\max\{p_A(n) : 0 \leq n \leq x\} = p_A(u).$$

15.2 Partition Functions for Finite Sets

Theorem 15.2 *Let A be a nonempty finite set of relatively prime positive integers, with $|A| = k$. Let $p_A(n)$ denote the number of partitions of n into parts belonging to A. Then*
$$p_A(n) = \left(\frac{1}{\prod_{a\in A} a}\right)\frac{n^{k-1}}{(k-1)!} + O\left(n^{k-2}\right).$$

Proof. The proof is by induction on k. If $k = 1$, then $A = \{1\}$ and $p_A(n) = 1$, since every positive integer has a unique partition into a sum of 1's.

Let $k \geq 2$, and assume that the theorem holds for $k-1$. Let $A = \{a_1, \ldots, a_k\}$. Then $\gcd(A) = (a_1, \ldots, a_k) = 1$. If $d = (a_1, \ldots, a_{k-1})$, then $(d, a_k) = 1$. For $i = 1, \ldots, k-1$ we set
$$a_i' = \frac{a_i}{d}.$$

Then $\gcd(a'_1, \ldots, a'_{k-1}) = 1$, and

$$A' = \{a'_1, \ldots, a'_{k-1}\}$$

is a set of $k-1$ relatively prime positive integers. Since the induction assumption holds for A', we have

$$p_{A'}(n) = \left(\frac{1}{\prod_{i=1}^{k-1} a'_i} \right) \frac{n^{k-2}}{(k-2)!} + O\left(n^{k-3}\right)$$

for all nonnegative integers n.

Let $n \geq (d-1)a_k$. Since $(d, a_k) = 1$, there exists a unique integer u such that $0 \leq u \leq d-1$ and

$$n \equiv ua_k \pmod{d}.$$

Then

$$m = \frac{n - ua_k}{d}$$

is a nonnegative integer, and

$$0 \leq m \leq n.$$

If v is any nonnegative integer such that

$$n \equiv va_k \pmod{d},$$

then $va_k \equiv ua_k \pmod{d}$, and so $v \equiv u \pmod{d}$, that is, $v = u + \ell d$ for some nonnegative integer ℓ. If

$$n - va_k = n - (u + \ell d)a_k \geq 0,$$

then

$$0 \leq \ell \leq \left[\frac{n}{da_k} - \frac{u}{d} \right] = \left[\frac{m}{a_k} \right] = r \leq m.$$

Let π be a partition of n into parts belonging to A. If π contains exactly v parts equal to a_k, then $n - va_k \geq 0$ and $n - va_k \equiv 0 \pmod{d}$, since $n - va_k$ is a sum of elements in $\{a_1, \ldots, a_{k-1}\}$ and each of the elements in this set is divisible by d. Therefore, $v = u + \ell d$, where $0 \leq \ell \leq r$. Consequently, we can divide the partitions of n with parts in A into $r+1$ classes, where, for each $\ell = 0, 1, \ldots, r$, a partition belongs to class ℓ if it contains exactly $u + \ell d$ parts equal to a_k. The number of partitions of n with exactly $u + \ell d$ parts equal to a_k is exactly the number of partitions of $n - (u + \ell d)a_k$ into parts belonging to the set $\{a_1, \ldots, a_{k-1}\}$, or, equivalently, the number of partitions of

$$\frac{n - (u + \ell d)a_k}{d} = m - \ell a_k$$

into parts belonging to A', which is exactly $p_{A'}(m-\ell a_k)$. Therefore,

$$\begin{aligned}
p_A(n) &= \sum_{\ell=0}^{r} p_{A'}(m-\ell a_k) \\
&= \left(\frac{1}{\prod_{i=1}^{k-1} a_i'}\right) \sum_{\ell=0}^{r} \left(\frac{(m-\ell a_k)^{k-2}}{(k-2)!} + O(m^{k-3})\right) \\
&= \left(\frac{d^{k-1}}{\prod_{i=1}^{k-1} a_i}\right) \sum_{\ell=0}^{r} \frac{(m-\ell a_k)^{k-2}}{(k-2)!} + O(n^{k-2}).
\end{aligned}$$

We evaluate the sum as follows. Since

$$\sum_{\ell=0}^{r} \ell^j = \frac{r^{j+1}}{(j+1)} + O(r^j)$$

by Exercise 5, and since

$$\sum_{j=0}^{k-2} (-1)^j \binom{k-1}{j+1} = -\sum_{j=1}^{k-1} (-1)^j \binom{k-1}{j} = 1,$$

we have

$$\begin{aligned}
&\sum_{\ell=0}^{r} \frac{(m-\ell a_k)^{k-2}}{(k-2)!} \\
&= \frac{1}{(k-2)!} \sum_{\ell=0}^{r} \sum_{j=0}^{k-2} \binom{k-2}{j} m^{k-2-j} (-\ell a_k)^j \\
&= \frac{1}{(k-2)!} \sum_{j=0}^{k-2} \binom{k-2}{j} m^{k-2-j} (-a_k)^j \sum_{\ell=0}^{r} \ell^j \\
&= \frac{1}{(k-2)!} \sum_{j=0}^{k-2} \binom{k-2}{j} m^{k-2-j} (-a_k)^j \left(\frac{r^{j+1}}{j+1} + O(r^j)\right) \\
&= \frac{1}{(k-2)!} \sum_{j=0}^{k-2} \binom{k-2}{j} m^{k-2-j} (-a_k)^j \left(\frac{m^{j+1}}{a_k^{j+1}(j+1)} + O(m^j)\right) \\
&= \frac{m^{k-1}}{a_k} \sum_{j=0}^{k-2} \binom{k-2}{j} \frac{(-1)^j}{(k-2)!(j+1)} + O(m^{k-2}) \\
&= \frac{m^{k-1}}{a_k} \sum_{j=0}^{k-2} \frac{(-1)^j}{(k-2-j)!j!(j+1)} + O(m^{k-2}) \\
&= \frac{m^{k-1}}{a_k} \sum_{j=0}^{k-2} \frac{(-1)^j}{(k-1-(j+1))!(j+1)!} + O(m^{k-2})
\end{aligned}$$

$$
\begin{aligned}
&= \frac{m^{k-1}}{a_k(k-1)!}\sum_{j=0}^{k-2}(-1)^j\binom{k-1}{j+1} + O(m^{k-2}) \\
&= \frac{m^{k-1}}{a_k(k-1)!} + O(m^{k-2}).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
p_A(n) &= \left(\frac{d^{k-1}}{\prod_{i=1}^{k-1} a_i}\right)\sum_{\ell=0}^{r}\frac{(m-\ell a_k)^{k-2}}{(k-2)!} + O(n^{k-2}) \\
&= \left(\frac{d^{k-1}}{\prod_{i=1}^{k-1} a_i}\right)\left(\frac{m^{k-1}}{a_k(k-1)!} + O(m^{k-2})\right) + O(n^{k-2}) \\
&= \left(\frac{1}{\prod_{i=1}^{k} a_i}\right)\frac{(n-ua_k)^{k-1}}{(k-1)!} + O(n^{k-2}) \\
&= \left(\frac{1}{\prod_{i=1}^{k} a_i}\right)\frac{n^{k-1}}{(k-1)!} + O(n^{k-2}).
\end{aligned}
$$

This completes the proof. □

Corollary 15.1 *Let $p_k(n)$ denote the number of partitions of n into at most k parts. Then*

$$p_k(n) \sim \frac{n^{k-1}}{k!(k-1)!} + O(n^{k-2}).$$

Proof. We know that $p_k(n)$ is also equal to the number of partitions of n into parts no greater than k. The result follows from Theorem 15.2 applied to the set $A = \{1, 2, \ldots, k\}$. □

Corollary 15.2 *Let A be an infinite set of positive integers with* $\gcd(A) = 1$. *Then*

$$\lim_{n\to\infty}\frac{\log p_A(n)}{\log n} = \infty.$$

Proof. For every sufficiently large integer k there exists a subset F_k of A of cardinality k such that $\gcd(F_k) = 1$. By Theorem 15.2,

$$p_A(n) \geq p_{F_k}(n) = \frac{n^{k-1}}{(k-1)!\prod_{a\in F_k} a} + O\left(n^{k-2}\right),$$

and so there exists a positive constant c_k such that

$$p_A(n) \geq c_k n^{k-1}$$

for all sufficiently large integers n. Then

$$\log p_A(n) \geq \log p_{F_k}(n) \geq (k-1)\log n + \log c_k.$$

Dividing by $\log n$, we obtain

$$\liminf_{n\to\infty} \frac{\log p_A(n)}{\log n} \geq k - 1.$$

This is true for all sufficiently large k, and so

$$\lim_{n\to\infty} \frac{\log p_A(n)}{\log n} = \infty.$$

This completes the proof. □

We can also use generating functions to compute partition functions of finite sets. For example, let $A = \{1, 2, 4\}$. By Theorem 15.2, we have

$$p_A(n) \sim \frac{n^2}{16} + O(n).$$

Using the partial fraction decomposition of the generating function, we can obtain an exact formula for $p_A(n)$ that is stronger than this asymptotic estimate. We have

$$\begin{aligned}
\sum_{n=0}^{\infty} p_A(n)x^n &= \frac{1}{(1-x)(1-x^2)(1-x^4)} \\
&= \frac{1}{(1-x)^3(1+x)^2(1+x^2)} \\
&= \frac{9}{32(1-x)} + \frac{1}{4(1-x)^2} + \frac{1}{8(1-x)^3} \\
&\quad + \frac{5}{32(1+x)} + \frac{1}{16(1+x)^2} + \frac{1+x}{8(1+x^2)}.
\end{aligned}$$

We write each partial fraction as a power series:

$$\begin{aligned}
\frac{9}{32(1-x)} &= \sum_{n=0}^{\infty} \frac{9}{32} x^n \\
\frac{1}{4(1-x)^2} &= \sum_{n=0}^{\infty} \frac{(n+1)}{4} x^n \\
\frac{1}{8(1-x)^3} &= \sum_{n=0}^{\infty} \frac{(n+2)(n+1)}{16} x^n \\
\frac{5}{32(1+x)} &= \sum_{n=0}^{\infty} \frac{(-1)^n 5}{32} x^n
\end{aligned}$$

$$\begin{aligned}
\frac{1}{16(1+x)^2} &= \sum_{n=0}^{\infty} \frac{(-1)^n(n+1)}{16} x^n \\
\frac{1+x}{8(1+x^2)} &= \sum_{n=0}^{\infty} \frac{(-1)^n(1+x)}{8} x^{2n} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{8} x^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n}{8} x^{2n+1} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^{[n/2]}}{8} x^n.
\end{aligned}$$

Therefore,

$$\begin{aligned}
p_A(n) &= \frac{9}{32} + \frac{n+1}{4} + \frac{(n+2)(n+1)}{16} + \frac{(-1)^n 5}{32} \\
&\quad + \frac{(-1)^n(n+1)}{16} + \frac{(-1)^{[n/2]}}{8} \\
&= \frac{n^2 + (7+(-1)^n)n}{16} + \frac{21 + (-1)^n 7 + (-1)^{[n/2]} 4}{32}.
\end{aligned}$$

If n is even, then

$$\begin{aligned}
p_A(n) &= \frac{n^2+8n+16}{16} + \frac{(-1)^{[n/2]} - 1}{8} \\
&= \begin{cases} \frac{(n+4)^2}{16} & \text{if } n \equiv 0 \pmod 4, \\ \frac{(n+4)^2}{16} - \frac{1}{4} & \text{if } n \equiv 2 \pmod 4. \end{cases}
\end{aligned}$$

If n is odd, then

$$\begin{aligned}
p_A(n) &= \frac{n^2+6n+9}{16} + \frac{(-1)^{[n/2]} - 1}{8} \\
&= \begin{cases} \frac{(n+3)^2}{16} & \text{if } n \equiv 1 \pmod 4, \\ \frac{(n+3)^2}{16} - \frac{1}{4} & \text{if } n \equiv 3 \pmod 4. \end{cases}
\end{aligned}$$

Exercises

1. Let $p_2(n)$ denote the number of partitions of n into at most 2 parts. Prove that
$$p_2(n) = \left[\frac{n}{2}\right] + 1.$$

2. Let $a \geq 2$ and $A = \{1, a\}$. Prove that
$$p_A(n) = \left[\frac{n}{a}\right] + 1.$$

3. Let $A = \{2, 3\}$. Prove that

$$p_A(n) = \begin{cases} \left[\frac{n}{6}\right] + 1 & \text{if } n \text{ is even and } n \geq 2, \\ \left[\frac{n-3}{6}\right] + 1 & \text{if } n \text{ is odd and } n \geq 3. \end{cases}$$

4. Let $A = \{2, a\}$, where a is an odd integer, $a \geq 3$. Compute $p_A(n)$.

5. Prove that

$$\sum_{\ell=0}^{r} \ell^j = \frac{r^{j+1}}{(j+1)} + O(r^j).$$

6. Let $A = \{1, 2, 3\}$. Let $\rho = (-1 + i\sqrt{3})/2$. Confirm the partial fraction decomposition

$$\begin{aligned} \sum_{n=1}^{\infty} p_A(n)x^n &= \frac{1}{(1-x)(1-x^2)(1-x^3)} \\ &= \frac{1}{(1-x)^3(1+x)(1-\rho x)(1-\rho^2 x)} \\ &= \frac{1}{6(1-x)^3} + \frac{1}{4(1-x)^2} + \frac{17}{72(1-x)} \\ &\quad + \frac{1}{8(1+x)} + \frac{1}{9(1-\rho x)} + \frac{1}{9(1-\rho^2 x)}. \end{aligned}$$

Show that this implies that

$$\begin{aligned} p_A(n) &= \frac{(n+2)(n+1)}{12} + \frac{n+1}{4} + \frac{17}{72} + \frac{(-1)^n}{8} + \frac{1}{9}(\rho^n + \rho^{2n}) \\ &= \frac{(n+3)^2}{12} - \frac{7}{72} + \frac{(-1)^n}{8} + \frac{1}{9}(\rho^n + \rho^{2n}) \\ &= \frac{(n+3)^2}{12} + r(n), \end{aligned}$$

where

$$|r(n)| < \frac{1}{2}.$$

Conclude that $p_A(n)$ is equal to the integer closest to $(n+3)^2/12$.

7. Let $p_k(n)$ denote the number of partitions of n into at most k parts. Show that the average number of parts in a partition of n is

$$\overline{p}(n) = \frac{1}{p(n)} \sum_{k=1}^{n} k\left(p_k(n) - p_{k-1}(n)\right).$$

Remark. Erdős and Lehner [35] proved that $\overline{p}(n) \sim c_0^{-1}\sqrt{n}\log n$.

15.3 Upper and Lower Bounds for $\log p(n)$

In this section we give Erdős's elementary proof that $\log p(n) \sim c_0\sqrt{n}$. We begin with some estimates for exponential functions.

Define $p(0) = 1$ and $p(-n) = 0$ for all $n \geq 1$.

Lemma 15.1 *If $0 < \ell \leq n$, then*

$$\sqrt{n} - \frac{\ell}{2\sqrt{n}} - \frac{\ell^2}{2n^{3/2}} \leq \sqrt{n-\ell} < \sqrt{n} - \frac{\ell}{2\sqrt{n}}.$$

Proof. If $0 < x \leq 1$, then

$$1 - \frac{x}{2} - \frac{x^2}{2} \leq (1-x)^{1/2} < 1 - \frac{x}{2}.$$

The result follows by letting $x = \ell/n$. □

Lemma 15.2 *If $x > 0$, then*

$$\frac{e^{-x}}{(1-e^{-x})^2} < \frac{1}{x^2}.$$

If $0 < x \leq 1$, then

$$\frac{e^{-x}}{(1-e^{-x})^2} > \frac{1}{x^2} - 2.$$

Proof. The power series expansion for e^x gives

$$\begin{aligned} e^{x/2} - e^{-x/2} &= 2\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\left(\frac{x}{2}\right)^{2k+1} \\ &= x + x^3 \sum_{k=1}^{\infty} \frac{x^{2k-2}}{(2k+1)!2^{2k}}. \end{aligned}$$

If $x > 0$, then

$$e^{x/2} - e^{-x/2} > x,$$

and so

$$\frac{e^{-x}}{(1-e^{-x})^2} = \frac{1}{\left(e^{x/2} - e^{-x/2}\right)^2} < \frac{1}{x^2}.$$

If $0 < x \leq 1$, then

$$e^{x/2} - e^{-x/2} < x + x^3 \sum_{k=1}^{\infty} \frac{1}{2^{2k}} < x + x^3 < \frac{x}{1-x^2},$$

and so

$$\frac{e^{-x}}{(1-e^{-x})^2} = \frac{1}{\left(e^{x/2}-e^{-x/2}\right)^2} > \left(\frac{1}{x} - x\right)^2 > \frac{1}{x^2} - 2.$$

□

Lemma 15.3 *Let c be a positive real number and let n be a positive integer. Then*

$$\sum_{k=1}^{\infty} \frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} < \frac{2\pi^2 n}{3c^2}.$$

If $n \geq c^2/4$, then

$$\sum_{k=1}^{\infty} \frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} > \frac{2\pi^2 n}{3c^2} - \frac{8\sqrt{n}}{c}.$$

Proof. Let k be a positive integer and

$$x = \frac{ck}{2\sqrt{n}}.$$

By Lemma 15.2,

$$\frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} = \frac{e^{-x}}{(1-e^{-x})^2} < \frac{1}{x^2} = \frac{4n}{c^2k^2},$$

and so

$$\sum_{k=1}^{\infty} \frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} < \frac{4n}{c^2}\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{4\pi^2 n}{6c^2} = \frac{2\pi^2 n}{3c^2}.$$

If $\sqrt{n} \geq c/2$ and $1 \leq k \leq 2\sqrt{n}/c$, then $0 < x \leq 1$ and, by Lemma 15.2,

$$\frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} > \frac{1}{x^2} - 2 = \frac{4n}{c^2k^2} - 2.$$

Therefore,

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} &> \sum_{k\leq 2\sqrt{n}/c} \frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} \\
&> \sum_{k\leq 2\sqrt{n}/c} \left(\frac{4n}{c^2k^2} - 2\right)
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{4n}{c^2}\left(\sum_{k=1}^{\infty}\frac{1}{k^2} - \sum_{k>2\sqrt{n}/c}\frac{1}{k^2}\right) - \frac{4\sqrt{n}}{c} \\
&= \frac{2\pi^2 n}{3c^2} - \frac{4n}{c^2}\sum_{k=[2\sqrt{n}/c]+1}^{\infty}\frac{1}{k^2} - \frac{4\sqrt{n}}{c}.
\end{aligned}
$$

For $k \geq 1$ we have

$$
\frac{1}{k^2} < \frac{1}{k^2 - 1/4} = \int_{k-1/2}^{k+1/2}\frac{dt}{t^2},
$$

and so

$$
\begin{aligned}
\frac{4n}{c^2}\sum_{k=[2\sqrt{n}/c]+1}^{\infty}\frac{1}{k^2} &< \frac{4n}{c^2}\int_{[2\sqrt{n}/c]+1/2}^{\infty}\frac{dt}{t^2} = \frac{4n}{c^2}\frac{1}{[2\sqrt{n}/c]+1/2} \\
&< \frac{4n}{c^2}\frac{1}{2\sqrt{n}/c - 1/2} \leq \frac{4\sqrt{n}}{c}.
\end{aligned}
$$

In the last inequality we used the fact that $\sqrt{n} \geq c/2$. Therefore,

$$
\sum_{k=1}^{\infty}\frac{e^{-\frac{ck}{2\sqrt{n}}}}{(1-e^{-\frac{ck}{2\sqrt{n}}})^2} > \frac{2\pi^2 n}{3c^2} - \frac{8\sqrt{n}}{c}.
$$

□

Lemma 15.4 *Let* $0 \leq t < 1$. *Then*

$$
\sum_{v=1}^{\infty} vt^v = \frac{t}{(1-t)^2}
$$

and

$$
\sum_{v=1}^{\infty} v^3 t^v = \frac{t^3 + 4t^2 + t}{(1-t)^4} \leq \frac{6t}{(1-t)^4}.
$$

Proof. Differentiating the power series

$$
\frac{1}{1-t} = \sum_{v=0}^{\infty} t^v,
$$

we obtain

$$
\frac{1}{(1-t)^2} = \sum_{v=1}^{\infty} vt^{v-1},
$$

$$\begin{aligned}
\frac{2}{(1-t)^3} &= \sum_{v=2}^{\infty} v(v-1)t^{v-2}, \\
\frac{6}{(1-t)^4} &= \sum_{v=3}^{\infty} v(v-1)(v-2)t^{v-3} \\
&= \sum_{v=3}^{\infty} (v^3 - 3v(v-1) - v)t^{v-3},
\end{aligned}$$

and so

$$\sum_{v=3}^{\infty} v^3 t^v = \frac{6t^3}{(1-t)^4} + 3t^2 \sum_{v=3}^{\infty} v(v-1)t^{v-2} + t\sum_{v=3}^{\infty} vt^{v-1}.$$

Then

$$\begin{aligned}
\sum_{v=1}^{\infty} v^3 t^v &= \frac{6t^3}{(1-t)^4} + 3t^2 \sum_{v=2}^{\infty} v(v-1)t^{v-2} + t\sum_{v=1}^{\infty} vt^{v-1} \\
&= \frac{6t^3}{(1-t)^4} + \frac{6t^2}{(1-t)^3} + \frac{t}{(1-t)^2} \\
&= \frac{t^3 + 4t^2 + t}{(1-t)^4} \\
&\leq \frac{6t}{(1-t)^4}.
\end{aligned}$$

□

Theorem 15.3

$$\log p(n) \sim c_0\sqrt{n}.$$

Proof. We shall use induction to obtain upper and lower bounds on $p(n)$. First we prove that

$$p(n) \leq e^{c_0\sqrt{n}} \tag{15.5}$$

for all nonnegative integers n. This is clearly true for $n = 0$ and $n = 1$. Let $n \geq 2$, and assume that the inequality holds for all integers strictly smaller than n. The notation $\sum_{kv \leq n}$ means the sum over all positive integers k and v such that $kv \leq n$. We have

$$\begin{aligned}
np(n) &= \sum_{kv\leq n} vp(n-kv) \leq \sum_{kv\leq n} ve^{c_0\sqrt{n-kv}} \\
&\leq \sum_{kv\leq n} ve^{c_0\sqrt{n} - \frac{c_0 kv}{2\sqrt{n}}} \qquad \text{(by Lemma 15.1)}
\end{aligned}$$

$$
\begin{aligned}
&\le e^{c_0\sqrt{n}} \sum_{k=1}^{\infty}\sum_{v=1}^{\infty} v\left(e^{-\frac{c_0 k}{2\sqrt{n}}}\right)^{v} \\
&= e^{c_0\sqrt{n}} \sum_{k=1}^{\infty} \frac{e^{-\frac{c_0 k}{2\sqrt{n}}}}{\left(1-e^{-\frac{c_0 k}{2\sqrt{n}}}\right)^2} \qquad \text{(by Lemma 15.4)} \\
&< \left(\frac{2\pi^2}{3c_0^2}\right) n e^{c_0\sqrt{n}} \qquad \text{(by Lemma 15.3)} \\
&= n e^{c_0\sqrt{n}}.
\end{aligned}
$$

This gives the upper bound (15.5).

Next we shall prove that for every ε with

$$0 < \varepsilon < c_0$$

there exists a constant $A = A(\varepsilon) > 0$ such that

$$p(n) \ge A e^{(c_0-\varepsilon)\sqrt{n}} \tag{15.6}$$

for all positive integers n. We begin by letting $A = e^{-c_0}$. Then (15.6) holds for $n = 1$, since $p(1) = 1 > e^{-\varepsilon} = A e^{c_0-\varepsilon}$.

Let $n \ge 2$, and assume that (15.6) holds for all integers less than n. Then

$$
\begin{aligned}
np(n) &= \sum_{kv\le n} v p(n-kv) \\
&\ge A \sum_{kv\le n} v e^{(c_0-\varepsilon)\sqrt{n-kv}} \\
&\ge A \sum_{kv\le n} v e^{(c_0-\varepsilon)\left(\sqrt{n}-\frac{kv}{2\sqrt{n}}-\frac{k^2v^2}{2n^{3/2}}\right)} \qquad \text{(by Lemma 15.1)} \\
&= A e^{(c_0-\varepsilon)\sqrt{n}} \sum_{kv\le n} v e^{-(c_0-\varepsilon)\left(\frac{kv}{2\sqrt{n}}+\frac{k^2v^2}{2n^{3/2}}\right)}.
\end{aligned}
$$

We shall show that

$$\sum_{kv\le n} v e^{-(c_0-\varepsilon)\left(\frac{kv}{2\sqrt{n}}+\frac{k^2v^2}{2n^{3/2}}\right)} \ge n.$$

Since $e^{-x} \ge 1 - x$, we have

$$e^{-\left((c_0-\varepsilon)\frac{k^2v^2}{2n^{3/2}}\right)} \ge 1 - \frac{(c_0-\varepsilon)k^2v^2}{2n^{3/2}},$$

and so

$$\sum_{kv\le n} v e^{-(c_0-\varepsilon)\left(\frac{kv}{2\sqrt{n}}+\frac{k^2v^2}{2n^{3/2}}\right)}$$

$$\begin{aligned} &\geq \sum_{kv\leq n} ve^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} - \frac{(c_0-\varepsilon)}{2n^{3/2}} \sum_{kv\leq n} k^2 v^3 e^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} \\ &= S_1(n) - \frac{(c_0-\varepsilon)}{2n^{3/2}} S_2(n). \end{aligned}$$

We shall estimate the sums $S_1(n)$ and $S_2(n)$.

If $kv > n$, then

$$\frac{(c_0-\varepsilon)kv}{2\sqrt{n}} > \frac{(c_0-\varepsilon)\sqrt{n}}{2} > \frac{(c_0-\varepsilon)}{2} > 0.$$

Since

$$e^{-t} \ll t^{-6} \qquad \text{for } t \geq (c_0-\varepsilon)/2,$$

we have

$$\begin{aligned} \sum_{kv>n} ve^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} &\ll \sum_{kv>n} v\left(\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}\right)^{-6} \\ &\ll n^3 \sum_{kv>n} \frac{1}{k^6 v^5} \\ &\ll n^3 \sum_{kv>n} \frac{1}{(kv)^{7/2} k^{5/2} v^{3/2}} \\ &< \frac{1}{\sqrt{n}} \sum_{k=1}^{\infty} \frac{1}{k^{5/2}} \sum_{v=1}^{\infty} \frac{1}{v^{3/2}} \\ &\ll \frac{1}{\sqrt{n}}. \end{aligned}$$

Then

$$\begin{aligned} S_1(n) &= \sum_{kv\leq n} ve^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} \\ &= \sum_{k=1}^{\infty}\sum_{v=1}^{\infty} ve^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} - \sum_{kv>n} ve^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} \\ &= \sum_{k=1}^{\infty} \frac{e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} + O\left(\frac{1}{\sqrt{n}}\right) \qquad \text{(by Lemma 15.4)} \\ &> \frac{2\pi^2 n}{3(c_0-\varepsilon)^2} + O(\sqrt{n}) \qquad \text{(by Lemma 15.3)} \\ &> \left(1+\frac{2\varepsilon}{c_0}\right) n + O(\sqrt{n}), \end{aligned}$$

since

$$\begin{aligned}\frac{2\pi^2}{3(c_0-\varepsilon)^2} &= \left(\frac{c_0}{c_0-\varepsilon}\right)^2 = \left(1+\frac{\varepsilon}{c_0-\varepsilon}\right)^2 \\ &> 1+\frac{2\varepsilon}{c_0-\varepsilon} > 1+\frac{2\varepsilon}{c_0}.\end{aligned}$$

We estimate the sum $S_2(n)$ as follows:

$$\begin{aligned}S_2(n) &= \sum_{kv\le n} k^2 v^3 e^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} \\ &\le \sum_{k=1}^{n} k^2 \sum_{v=1}^{\infty} v^3 e^{-\frac{(c_0-\varepsilon)kv}{2\sqrt{n}}} \\ &\le 6\sum_{k=1}^{n} \frac{k^2 e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^4} \qquad \text{(by Lemma 15.4)} \\ &= 6\sum_{k=1}^{n} \frac{e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} \frac{k^2}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} \\ &< 6\sum_{k=1}^{n} \left(\frac{4n}{(c_0-\varepsilon)^2k^2}\right) \frac{k^2}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} \qquad \text{(by Lemma 15.2)} \\ &\ll n\sum_{k=1}^{n} \frac{1}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2}.\end{aligned}$$

Let

$$x = \frac{(c_0-\varepsilon)k}{2\sqrt{n}}.$$

If $1 \le k \le \sqrt{n}$, then $0 < x < c_0/2$ and

$$1-e^{-x} = \int_0^x e^{-t}dt \ge xe^{-x} > xe^{-c_0/2},$$

and so

$$\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2 = \left(1-e^{-x}\right)^2 > x^2e^{-c_0} = \frac{e^{-c_0}(c_0-\varepsilon)^2k^2}{4n}.$$

Therefore,

$$\sum_{1\le k\le\sqrt{n}} \frac{1}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} < \frac{4e^{c_0}n}{(c_0-\varepsilon)^2} \sum_{1\le k\le\sqrt{n}} \frac{1}{k^2} \ll n.$$

If $k > \sqrt{n}$, then

$$\sum_{\sqrt{n}<k\leq n} \frac{1}{\left(1-e^{-\frac{(c_0-\varepsilon)k}{2\sqrt{n}}}\right)^2} < \sum_{\sqrt{n}<k\leq n} \frac{1}{\left(1-e^{-\frac{(c_0-\varepsilon)}{2}}\right)^2} \ll n.$$

Therefore,

$$S_2(n) \ll n^2.$$

Since

$$S_1(n) > 0 \quad \text{and} \quad S_2(n) > 0,$$

we have

$$\begin{aligned} S_1(n) - \frac{(c_0-\varepsilon)}{2n^{3/2}} S_2(n) &\geq \left(1+\frac{2\varepsilon}{c_0}\right) n + O(\sqrt{n}) - \frac{(c_0-\varepsilon)}{2n^{3/2}} O(n^2) \\ &> \left(1+\frac{2\varepsilon}{c_0}\right) n - c_1\sqrt{n} \end{aligned}$$

for some positive constant c_1. Then

$$\begin{aligned} np(n) &\geq Ae^{(c_0-\varepsilon)\sqrt{n}} \left(S_1(n) - \frac{(c_0-\varepsilon)k}{2n^{3/2}} S_2(n)\right) \\ &\geq Ane^{(c_0-\varepsilon)\sqrt{n}} + A\sqrt{n}e^{(c_0-\varepsilon)\sqrt{n}} \left(\frac{2\varepsilon\sqrt{n}}{c_0} - c_1\right) \\ &> Ane^{(c_0-\varepsilon)\sqrt{n}} \end{aligned}$$

if we choose $A > 0$ small enough that (15.6) holds for all $n \leq (c_0c_1/2\varepsilon)^2$.

It follows from (15.5) and (15.6) that for every $\varepsilon > 0$ there exists a constant A such that

$$(c_0-\varepsilon)\sqrt{n} + \log A < \log p(n) < c_0\sqrt{n}$$

for all positive integers n, and so $\log p(n) \sim c_0\sqrt{n}$. This completes the proof of the theorem. □

Exercises

1. Prove that the recursion formula (15.3) is equivalent to

$$np(n) = \sum_{\nu=1}^{\infty} \sigma(\nu)p(n-\nu).$$

15.4 Notes

In 1918 Hardy and Ramanujan [59, 58] published the asymptotic formula for the partition function. Uspensky [146] obtained the same result independently in 1920. Both papers used complex variables and modular functions to deduce the asymptotic estimate $p(n) \sim (4n\sqrt{3})^{-1}e^{c_0\sqrt{n}}$. In their 1918 paper, Hardy and Ramanujan wrote,

> it is equally possible to prove [$\log p(n) \sim c_0\sqrt{n}$] by reasoning of a more elementary, though more special character; we have a proof, for example, based on the identity
>
> $$np(n) = \sum_{\nu=1}^{\infty} \sigma(\nu)p(n-\nu), \tag{15.7}$$
>
> where $\sigma(\nu)$ is the sum of the divisors of ν, and a process of induction.

Many years later, however, Hardy wrote in his book *Ramanujan* [52, p. 114],

> It is actually true that $\log p(n) \sim \pi\sqrt{(2n/3)}\ldots$, but we cannot prove this very simply.

Hardy and Ramanujan clearly had no elementary proof of the asymptotic formula (15.1); in their 1918 paper they wrote that

> we are at present unable to obtain, by any method which does not depend upon Cauchy's theorem, a result as precise as [$p(n) \sim e^{c_0\sqrt{n}}/(4\sqrt{3})n$], a result, that is to say, which is "vraiment asymptotique."

Erdős's proof of the asymptotic formula for $p(n)$, published in 1942 in [32], is a *tour de force* of elementary methods in number theory. This proof is not as famous nor as controversial as the elementary proof of the prime number theorem, but it is impressive in its depth and technical difficulty. It shows that the asymptotic formula for $p(n)$ is simply a consequence of the elementary recursion formula (15.7), and is independent of any deep analytic properties of modular functions.

Knessl and Keller [80] develop Erdős's method and apply the recursion formula for the partition function to derive formal asymptotic expansions.

Grosswald [42] and Hua [68] have presented Erdős's elementary proof of (15.2). There is a different elementary proof of the upper bound $\log p(n) < \pi\sqrt{2n/3}$ in unpublished lectures of Siegel on analytic number theory; Siegel's proof appears in Knopp [81, pp. 88–90]. Analytic proofs of (15.1) can be found in Apostol [4], Knopp [81], and Rademacher [119].

The standard proof of Theorem 15.2 uses the partial fraction decomposition of a generating function. The proof in this book is due to Nathanson [107].

Let $P_k(n) = p_k(n) - p_{k-1}(n)$ denote the number of partitions of n into *exactly* k parts. Erdős [33] proved that for fixed n, the maximum value of $P_k(n)$ occurs for $k_0 \sim c_0^{-1} n^{1/2} \log n$. This had been conjectured by Auluck, Chowla, and Gupta [6]. Using hard analysis, Szekeres [137, 138] proved that for sufficiently large n, the finite sequence $P_k(n)$ is *unimodal* in the sense that there exists an integer k_0 such that $P_{k-1}(n) \leq P_k(n)$ for $1 \leq k \leq k_0$ and $P_{k-1}(n) \geq P_k(n)$ for $k_0 + 1 \leq k \leq n$. It would be very interesting to have an elementary proof of the unimodality of the partition function $P_k(n)$.

Rademacher [117, 118] obtained a convergent series for $p(n)$ of the form

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{k=1}^{\infty} k^{1/2} A_k(n) \frac{d}{dn} \frac{\sinh\left(\frac{\pi\lambda_n}{k}\sqrt{\frac{2}{3}}\right)}{\lambda_n}.$$

After studying the original paper of Hardy and Ramanujan, Selberg (unpublished) independently proved the same formula. Many years later he wrote [130], "I am inclined to believe that Rademacher and I were the only ones to have studied this paper thoroughly since the time it was written."

16

An Inverse Theorem for Partitions

16.1 Density Determines Asymptotics

Let A be a set of integers, and let $A(x)$ denote the number of positive elements of A that do not exceed x. Recall that $A(x)$ is called the *counting function* of A. Then $0 \leq A(x) \leq x$, and so $0 \leq A(x)/x \leq 1$ for all x. The set A has *asymptotic density* α if

$$\lim_{x\to\infty} \frac{A(x)}{x} = \alpha.$$

For example, the set of all positive integers has density 1, and every finite set has density 0. The set of even integers has density 1/2. By Chebyshev's theorem (Theorem 8.2), the set of prime numbers has density 0.

If A has density α, then for every $\varepsilon > 0$ there exists a number $x_0(\varepsilon)$ such that for all $x \geq x_0(\varepsilon)$,

$$\left| \frac{A(x)}{x} - \alpha \right| < \varepsilon,$$

or, equivalently,

$$(\alpha - \varepsilon)x < A(x) < (\alpha + \varepsilon)x. \tag{16.1}$$

There exists an integer $k_0(\varepsilon)$ such that if $a_k \in A$ and $k \geq k_0(\varepsilon)$, then $a_k \geq x_0(\varepsilon)$. Setting $x = a_k$ in inequality (16.1), we obtain

$$(\alpha - \varepsilon)a_k < k < (\alpha + \varepsilon)a_k,$$

and so

$$\frac{k}{\alpha + \varepsilon} < a_k < \frac{k}{\alpha - \varepsilon}.$$

In Chapter 15 we proved that $\log p(n) \sim c_0\sqrt{n}$. In this section we shall prove that if A is any set of integers of density $\alpha > 0$ and $\gcd(A) = 1$, then

$$\log p_A(n) \sim c_0\sqrt{\alpha n}. \tag{16.2}$$

In Section 16.2 we prove the converse: If A is any set of positive integers whose partition function $p_A(n)$ satisfies (16.2) for some $\alpha > 0$, then A has asymptotic density α.

A set of positive integers is *cofinite* if it contains all but finitely many positive integers. We begin with a simple result about partition functions of cofinite sets.

Lemma 16.1 *Let A be a cofinite set of positive integers. Then*

$$\log p_A(n) \sim c_0\sqrt{n}.$$

Proof. If A is cofinite, then A contains all sufficiently large integers. Choose a positive integer $\ell > 1$ such that A contains all integers greater than ℓ, that is,

$$B = \{n \geq \ell + 1\} \subseteq A.$$

Then

$$p_B(n) \leq p_A(n) \leq p(n).$$

Since $\log p(n) \sim c_0\sqrt{n}$, it suffices to prove that $\log p_B(n) \sim c_0\sqrt{n}$.

Consider the finite set $F = \{1, 2, \ldots, \ell\}$. Since $\gcd(F) = 1$, Theorem 15.2 implies that there exists a constant $c \geq 1$ such that $p_F(n) \leq cn^{\ell-1}$ for all positive integers n. Each part of an unrestricted partition of n belongs to F or to B, and so every partition of n is uniquely of the form $n = (n-m)+m$, where $n - m$ is a sum of elements of F and m is a sum of elements of B. By Exercise 4, the partition function $p_B(n)$ is increasing for $n \geq 1$, and so

$$\begin{aligned} p(n) &= \sum_{m=0}^{n} p_F(n-m)p_B(m) \\ &\leq cn^{\ell-1}\sum_{m=0}^{n} p_B(m) \\ &\leq 2cn^{\ell}p_B(n) \\ &\leq 2cn^{\ell}p(n). \end{aligned}$$

Taking logarithms and dividing by $c_0\sqrt{n}$, we have

$$\begin{aligned} \frac{\log p(n)}{c_0\sqrt{n}} &\leq \frac{\log 2c + \ell\log n}{c_0\sqrt{n}} + \frac{\log p_B(n)}{c_0\sqrt{n}} \\ &\leq \frac{\log 2c + (\ell-1)\log n}{c_0\sqrt{n}} + \frac{\log p(n)}{c_0\sqrt{n}}. \end{aligned}$$

Letting n go to infinity, we obtain $\log p_B(n) \sim c_0\sqrt{n}$. This completes the proof.

Theorem 16.1 *Let A be a set of positive integers. If A has density $\alpha > 0$ and* $\gcd(A) = 1$, *then the partition function $p_A(n)$ satisfies the asymptotic equation*

$$\log p_A(n) \sim c_0\sqrt{\alpha n}.$$

Proof. Let $A = \{a_k\}_{k=1}^{\infty}$, where $a_1 < a_2 < \cdots$. Let $0 < \varepsilon < \alpha$. Since $d(A) = \alpha$ and $\gcd(A) = 1$, there exists an integer $\ell_0 = \ell_0(\varepsilon)$ such that $\gcd\{a_k : 1 \leq k \leq \ell_0\} = 1$ and

$$\frac{k}{\alpha+\varepsilon} < a_k < \frac{k}{\alpha-\varepsilon} \tag{16.3}$$

for all $k > \ell_0$.

We begin by deriving the upper bound

$$\limsup_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \leq 1.$$

Let $F = \{a_1, a_2, \ldots, a_{\ell_0}\}$ and $B = \{a_k \in A : k \geq \ell_0 + 1\}$. Let m be a positive integer, $m \leq n$, and let

$$m = a_{k_1} + a_{k_2} + \cdots + a_{k_r}$$

be a partition of m with parts in B. To this partition of m we associate the partition

$$n' = k_1 + k_2 + \cdots + k_r.$$

By (16.3) we have $k_i < (\alpha + \varepsilon)a_{k_i}$, and so

$$\begin{aligned} n' &< (\alpha+\varepsilon)a_{k_1} + (\alpha+\varepsilon)a_{k_2} + \cdots + (\alpha+\varepsilon)a_{k_r} \\ &= (\alpha+\varepsilon)m \\ &\leq (\alpha+\varepsilon)n. \end{aligned}$$

This establishes a one-to-one mapping from partitions of m with parts in B to partitions of integers n' less than $(\alpha + \varepsilon)n$. Since the unrestricted partition function $p(n)$ is increasing, we have

$$\begin{aligned} p_B(m) &\leq \sum_{1\leq n'\leq(\alpha+\varepsilon)n} p(n') \\ &\leq (\alpha+\varepsilon)np([(\alpha+\varepsilon)n]) \\ &< 2np([(\alpha+\varepsilon)n]). \end{aligned}$$

Recall that $A = F \cup B$, where F consists of ℓ_0 relatively prime positive integers. By Theorem 15.2, there exists a constant c such that

$$p_F(n) \leq cn^{\ell_0-1}$$

for every positive integer n. Every partition of n with parts in A decomposes uniquely into a partition of m with parts in B and a partition of $n-m$ with parts in F for some nonnegative integer $m \leq n$. Then

$$\begin{aligned} p_A(n) &= \sum_{m=0}^{n} p_F(n-m)p_B(m) \\ &\leq cn^{\ell_0-1}\sum_{m=0}^{n} p_B(m) \\ &\leq cn^{\ell_0-1}\sum_{m=0}^{n} 2np([(\alpha+\varepsilon)n]) \\ &\leq 4cn^{\ell_0+1}p([(\alpha+\varepsilon)n]). \end{aligned}$$

Since $\log p(n) \sim c_0\sqrt{n}$, it follows that for every $\varepsilon > 0$ there exists an integer $n_0(\varepsilon)$ such that

$$\log p([(\alpha+\varepsilon)n]) < (1+\varepsilon)c_0\sqrt{[(\alpha+\varepsilon)n]}$$

for $n \geq n_0(\varepsilon)$. Therefore,

$$\begin{aligned} \log p_A(n) &\leq \log 4c + (\ell_0+1)\log n + \log p([(\alpha+\varepsilon)n]) \\ &< \log 4c + (\ell_0+1)\log n + (1+\varepsilon)c_0\sqrt{(\alpha+\varepsilon)n} \end{aligned}$$

for $n \geq \ell_0(\varepsilon)$. Dividing by $c_0\sqrt{\alpha n}$, we obtain

$$\frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \leq \frac{\log 4c + k_0 \log n}{c_0\sqrt{\alpha n}} + (1+\varepsilon)\sqrt{1+\frac{\varepsilon}{\alpha}},$$

and so

$$\limsup_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \leq (1+\varepsilon)\sqrt{1+\frac{\varepsilon}{\alpha}}.$$

This inequality is true for all $\varepsilon > 0$, and so

$$\limsup_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \leq 1.$$

Next we obtain the lower bound

$$\liminf_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \geq 1.$$

Since $\gcd(A) = 1$, Theorem 1.16 implies that $p_A(n) \geq 1$ for all sufficiently large n. For $0 < \varepsilon < \alpha$, there exists a positive integer $\ell_0 = \ell_0(\varepsilon)$ such that $\gcd\{a_k : 1 \leq k \leq \ell_0\} = 1$ and

$$\frac{k}{\alpha+\varepsilon} < a_k < \frac{k}{\alpha-\varepsilon}$$

for all $k > \ell_0$.

Let $p'(n)$ denote the number of partitions of n into parts greater than ℓ_0. To every partition

$$n = k_1 + \cdots + k_r \qquad \text{with } k_1 \geq \cdots \geq k_r > \ell_0,$$

we associate the partition

$$m = a_{k_1} + \cdots + a_{k_r}.$$

Inequality (16.3) implies that

$$m < \frac{n}{\alpha - \varepsilon}.$$

This is a one-to-one mapping from partitions of n with parts greater than ℓ_0 to partitions of integers $m < n/(\alpha - \varepsilon)$ with parts in A. Therefore,

$$\begin{aligned} p'(n) &\leq \sum_{m < \frac{n}{\alpha-\varepsilon}} p_A(m) \\ &< \frac{n}{\alpha - \varepsilon} \max\left\{ p_A(m) : m \leq \frac{n}{\alpha - \varepsilon} \right\} \\ &\leq \frac{n p_A(u_n)}{\alpha - \varepsilon}, \end{aligned}$$

where, by Exercise 7 of Section 15.1, u_n is an integer in the bounded interval

$$\frac{n}{\alpha - \varepsilon} - a_1 < u_n \leq \frac{n}{\alpha - \varepsilon}.$$

The sequence $\{u_n\}_{n=1}^{\infty}$ is not necessarily increasing, but

$$\lim_{n \to \infty} u_n = \infty.$$

Let d be the unique positive integer such that

$$0 < (\alpha - \varepsilon)a_1 \leq d < (\alpha - \varepsilon)a_1 + 1.$$

For every $i, j \geq 1$,

$$u_{(i+j)d} - u_{id} > \left(\frac{(i+j)d}{\alpha - \varepsilon} - a_1 \right) - \frac{id}{\alpha - \varepsilon} = \frac{jd}{\alpha - \varepsilon} - a_1 \geq (j-1)a_1.$$

It follows that $u_{(i+1)d} > u_{id}$, and so the sequence $\{u_{id}\}_{i=1}^{\infty}$ is strictly increasing. Similarly,

$$u_{(i+j)d} - u_{id} < \frac{(i+j)d}{\alpha - \varepsilon} - \left(\frac{id}{\alpha - \varepsilon} - a_1 \right) = \frac{jd}{\alpha - \varepsilon} + a_1 < (j+1)a_1 + \frac{j}{\alpha - \varepsilon}.$$

Choose N_0 such that $p_A(n) \geq N_0$ for all $n \geq N_0$. Let i_0 be the unique integer such that

$$\frac{N_0}{a_1} + 1 \leq i_0 < \frac{N_0}{a_1} + 2.$$

Then

$$u_{id} - u_{(i-i_0)d} > (i_0 - 1)a_1 \geq N_0$$

for all $i \geq i_0$. For every integer $n \geq i_0 d$ there exists an integer $i \geq i_0$ such that

$$u_{id} \leq n < u_{(i+1)d}.$$

Then

$$n - u_{(i-i_0)d} < u_{(i+1)d} - u_{(i-i_0)d} < (i_0 + 2)d + \frac{i_0 + 1}{\alpha - \varepsilon}$$

and

$$n - u_{(i-i_0)d} \geq u_{id} - u_{(i-i_0)d} > N_0.$$

Therefore,

$$p_A(n - u_{(i-i_0)d}) \geq 1.$$

By Exercise 6 of Section 15.1,

$$p_A(n) \geq p_A(u_{(i-i_0)d}) > \frac{(\alpha - \varepsilon)p'((i - i_0)d)}{(i - i_0)d}.$$

Since

$$n < u_{(i+1)d} \leq \frac{(i + 1)d}{\alpha - \varepsilon},$$

it follows that

$$(i - i_0)d > (\alpha - \varepsilon)n - (i_0 + 1)d$$

and

$$p_A(n) > \frac{(\alpha - \varepsilon)p'((\alpha - \varepsilon)n - (i_0 + 1)d)}{(i - i_0)d}.$$

Since $p'(n)$ is the partition function of a cofinite subset of the positive integers, Lemma 16.1 implies that for n sufficiently large,

$$\begin{aligned}
\log p_A(n) &> \log p'((\alpha - \varepsilon)n - (i_0 + 1)d)) + \log(\alpha - \varepsilon) - \log(i - i_0)d \\
&> (1 - \varepsilon)c_0\sqrt{(\alpha - \varepsilon)n - (i_0 + 1)d} + \log(\alpha - \varepsilon) - \log(i - i_0)d.
\end{aligned}$$

Dividing by $c_0\sqrt{\alpha n}$, we obtain

$$\liminf_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \geq (1 - \varepsilon)\sqrt{1 - \varepsilon/\alpha}.$$

This inequality holds for $0 < \varepsilon < \alpha$, and so

$$\liminf_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{\alpha n}} \geq 1.$$

This completes the proof.

Exercises

1. Prove that the set $\{2^k : k \geq 0\}$ has density 0. Prove that the set $\{2^k 3^\ell : k, \ell \geq 0\}$ has density 0.

2. Let A be a set of positive integers, and let $B = \mathbf{N} \setminus A$ be the set of positive integers not in A. Prove that if $d(A) = \alpha$, then $d(B) = 1 - \alpha$.

3. In this exercise we construct a set A that does not have a density. We denote by $(x, y]$ the set of integers n such that $x < n \leq y$. Let $N_1 < N_2 < N_3 < \cdots$ be a strictly increasing sequence of positive integers such that $\lim_{r\to\infty} N_{r+1}/N_r = \infty$, and let
$$A = \bigcup_{r=1}^{\infty} (N_{2r-1}, N_{2r}].$$
Prove that
$$\lim_{r\to\infty} \frac{A(N_{2r})}{N_{2r}} = 1$$
and
$$\lim_{r\to\infty} \frac{A(N_{2r+1})}{N_{2r+1}} = 0.$$
Since $\limsup_{x\to\infty} A(x)/x = 1$ and $\liminf_{x\to\infty} A(x)/x = 0$, the set A does not have an asymptotic density.

   Hint: Show that $A(N_{2r}) \geq N_{2r} - N_{2r-1}$ and $A(N_{2r+1}) \leq N_{2r}$.

4. We say that a partition $a_1 + a_2 + \cdots + a_r$ has a *unique largest part* if $a_1 > a_2 \geq \cdots \geq a_r$. Let n_0 be a positive integer, and let A be the set of all integers greater than or equal to n_0. Show that $p_A(n) = 1$ for $n_0 \leq n < 2n_0$. Let $n \geq n_0$. To every partition π of n we can associate a partition of $n + 1$ by adding 1 to the largest part of π. Show that this map is a bijection between partitions of n and partitions of $n+1$ with a unique largest part. Deduce that $p_A(n)$ is increasing for $n \geq 1$, and strictly increasing for sufficiently large n.

5. Let $a_1, \ldots, a_\ell$, and m be integers such that
$$1 \leq a_1 < \cdots < a_\ell \leq m$$
and
$$(a_1, \ldots, a_\ell, m) = 1.$$
Let A be the set of all positive integers a such that $a \equiv a_i \pmod{m}$ for some $i = 1, \ldots, \ell$. Prove that
$$\log p_A(n) \sim c_0 \sqrt{\frac{\ell n}{m}}.$$

6. Prove that if the set A of positive integers has positive density, then

$$d(A) = \lim_{n\to\infty} \left(\frac{\log p_A(n)}{\log p(n)} \right)^2 .$$

7. Let A be a set of positive integers. The *upper asymptotic density* of A is

$$d_U(A) = \limsup_{n\to\infty} \frac{A(n)}{n}.$$

Prove that if $\gcd(A) = 1$ and $d_U(A) \leq \alpha$, then

$$\limsup_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{n}} \leq \sqrt{\alpha}.$$

8. Let A be a set of positive integers. The *lower asymptotic density* of A is

$$d_L(A) = \liminf_{n\to\infty} \frac{A(n)}{n}.$$

Prove that if $\gcd(A) = 1$ and $d_L(A) \geq \alpha$, then

$$\liminf_{n\to\infty} \frac{\log p_A(n)}{c_0\sqrt{n}} \geq \sqrt{\alpha}.$$

9. Let A be a set of positive integers with $\gcd(A) = 1$. Prove that if $d(A) = 0$, then $\log p_A(n) = o(\sqrt{n})$.

16.2 Asymptotics Determine Density

The goal of this section is an inverse theorem for partitions. We shall prove that the asymptotics of the partition function $p_A(n)$ determines the density of the set A.

We begin with some remarks about generating functions. If a is a positive integer and $|x| < 1$, then the geometric progression

$$(1 - x^a)^{-1} = 1 + x^a + x^{2a} + x^{3a} + \cdots$$

converges absolutely. If A is a finite set of positive integers, then

$$\begin{aligned} \prod_{a\in A} (1 - x^a)^{-1} &= \prod_{a\in A} \left(1 + x^a + x^{2a} + x^{3a} + \cdots\right) \\ &= \sum_{n=0}^{\infty} p_A(n) x^n, \end{aligned}$$

where $p_A(n)$ is the partition function for A.

If A is an infinite set of positive integers and $|x| < 1$, then the infinite product

$$\prod_{a\in A}(1-x^a)^{-1}$$

converges absolutely, since

$$\sum_{a\in A}|x|^a \leq \sum_{a=1}^{\infty}|x|^a = \frac{|x|}{1-|x|} < \infty$$

and

$$f(x) = \prod_{a\in A}(1-x^a)^{-1} = \sum_{n=0}^{\infty}p_A(n)x^n.$$

This function is called the *generating function* for the partition function $p_A(n)$.

Theorem 16.2 *Let A be a set of positive integers with* $\gcd(A) = 1$. *Let $p_A(n)$ denote the number of partitions of n with parts in A. If there exists a number $\alpha > 0$ such that*

$$\log p_A(n) \sim c_0\sqrt{\alpha n},$$

then the set A has density α.

Proof. The proof uses an Abelian theorem (Theorem 16.3) and a Tauberian theorem (Theorem 16.4) that we prove in the next section. The generating function

$$f(x) = \sum_{n=1}^{\infty}p_A(n)x^n = \prod_{a\in A}(1-x^a)^{-1}$$

converges for $|x| < 1$. Since

$$\log p_A(n) \sim c_0\sqrt{\alpha n} = 2\sqrt{\frac{\pi^2\alpha n}{6}},$$

Theorem 16.3 immediately implies that

$$\log f(x) \sim \frac{\pi^2\alpha}{6(1-x)}.$$

Applying the Taylor series

$$-\log(1-x) = \sum_{k=1}^{\infty}\frac{x^k}{k}$$

for $|x| < 1$, we have

$$\log f(x) = -\sum_{a\in A} \log(1-x^a) = \sum_{a\in A}\sum_{k=1}^{\infty} \frac{x^{ak}}{k} = \sum_{n=1}^{\infty} b_n x^n,$$

where

$$b_n = \sum_{\substack{a\in A\\ n=ak}} \frac{1}{k} = \sum_{\substack{a\in A\\ a|n}} \frac{a}{n} \geq 0.$$

By Theorem 16.4,

$$S_B(x) = \sum_{n\leq x} b_n \sim \frac{\pi^2 \alpha x}{6}.$$

We define the remainder function $r(x)$ by

$$S_B(x) = \frac{\pi^2\alpha x}{6}(1+r(x)).$$

The function $S_B(x)$ is an increasing, nonnegative function such that $S_B(x) = 0$ for $x < 1$ and

$$\begin{aligned} S_B(x) &= \sum_{n\leq x}\sum_{\substack{a\in A\\ n=ak}} \frac{1}{k} \\ &= \sum_{k\leq x} \frac{1}{k} \sum_{\substack{a\in A\\ ak\leq x}} 1 \\ &= \sum_{k\leq x} \frac{1}{k} A\left(\frac{x}{k}\right), \end{aligned}$$

where $A(x)$ is the counting function of the set A. By Möbius inversion (Exercise 7 in Section 6.3), we have

$$A(x) = \sum_{k\leq x} \frac{\mu(k)}{k} S_B\left(\frac{x}{k}\right).$$

For every $\varepsilon > 0$ there exists a number $x_0 = x_0(\varepsilon)$ such that the remainder function $r(x)$ satisfies the inequality $|r(x)| < \varepsilon$ for all $x \geq x_0$. If $k \leq x/x_0$, then $x/k \geq x_0$ and $|r(x/k)| < \varepsilon$. If $k > x/x_0$, then $x/k < x_0$ and $0 \leq S_B(x/k) \leq S_B(x_0)$. Therefore,

$$\begin{aligned} A(x) &= \sum_{k\leq x} \frac{\mu(k)}{k} S_B\left(\frac{x}{k}\right) \\ &= \sum_{k\leq x/x_0} \frac{\mu(k)}{k} \left(\frac{\pi^2\alpha x}{6k}\left(1 + r\left(\frac{x}{k}\right)\right)\right) + \sum_{x/x_0<k\leq x} \frac{\mu(k)}{k} S_B\left(\frac{x}{k}\right) \end{aligned}$$

$$= \frac{\pi^2 \alpha x}{6} \sum_{k \leq x/x_0} \frac{\mu(k)}{k^2} + \frac{\pi^2 \alpha x}{6} \sum_{k \leq x/x_0} \frac{\mu(k)}{k^2} r\left(\frac{x}{k}\right) + \sum_{x/x_0 < k \leq x} \frac{\mu(k)}{k} S_B\left(\frac{x}{k}\right).$$

We estimate these three terms separately. By Theorem 6.17,

$$\sum_{k \leq x/x_0} \frac{\mu(k)}{k^2} = \frac{6}{\pi^2} - \sum_{k > x/x_0} \frac{\mu(k)}{k^2} = \frac{6}{\pi^2} + O\left(\frac{x_0}{x}\right),$$

and so

$$\frac{\pi^2 \alpha x}{6} \sum_{k \leq x/x_0} \frac{\mu(k)}{k^2} = \alpha x + O\left(x_0\right).$$

Similarly,

$$\left| \frac{\pi^2 \alpha x}{6} \sum_{k \leq x/x_0} \frac{\mu(k)}{k^2} r\left(\frac{x}{k}\right) \right| \leq \frac{\pi^2 \alpha \varepsilon x}{6} \sum_{k \leq x/x_0} \frac{1}{k^2} = O(\varepsilon x).$$

The third term is bounded independently of x, since

$$\begin{aligned} \left| \sum_{x/x_0 < k \leq x} \frac{\mu(k)}{k} S_B\left(\frac{x}{k}\right) \right| &\leq S_B(x_0) \sum_{x/x_0 < k \leq x} \frac{1}{k} \\ &\leq 2 S_B(x_0) \log x_0 \\ &= O(x_0). \end{aligned}$$

Therefore,

$$A(x) = \alpha x + O(\varepsilon x) + O(x_0) \sim \alpha x.$$

This completes the proof.

Exercises

We can use the Taylor series for the generating function for the unrestricted partition function $p(n)$ to obtain a simple proof of the upper bound $\log p(n) < c_0 \sqrt{n}$.

1. For $0 < x < 1$, let

$$f(x) = \prod_{n=1}^{\infty} (1 - x^n)^{-1} = \sum_{n=0}^{\infty} p(n) x^n.$$

   Prove that

$$\log p(n) + n \log x < \log f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k(1 - x^k)}.$$

2. Prove that if $0 < x < 1$, then
$$1 - x^k > kx^{k-1}(1-x)$$
and
$$\log f(x) < \frac{\pi^2 x}{6(1-x)}.$$

3. Prove that if $0 < x < 1$, then
$$-\log x < \frac{1-x}{x},$$
and so
$$\log p(n) < \frac{\pi^2 x}{6(1-x)} + \frac{n(1-x)}{x}.$$

4. Prove that $\log p(n) < c_0\sqrt{n}$.

   Hint: Choose $x \in (0,1)$ such that
$$\frac{\pi^2 x}{6(1-x)} = \frac{n(1-x)}{x}.$$

16.3 Abelian and Tauberian Theorems

In this section we derive the two results about power series with nonnegative coefficients that were used to deduce Theorem 16.2. The proofs require only advanced calculus. To the sequence $B = \{b_n\}_{n=0}^{\infty}$ of real numbers we can associate the power series $f(x) = \sum_{n=0}^{\infty} b_n x^n$. We shall assume that the power series converges for $|x| < 1$. We think of the function $f(x)$ as a kind of average over the sequence B. In rough language, an *Abelian theorem* asserts that if the sequence B has some property, then the function $f(x)$ has some related property. Conversely, a *Tauberian theorem* asserts that if the function $f(x)$ has some property, then the sequence B has a related property.

The following result is an Abelian theorem.

Theorem 16.3 *Let $B = \{b_n\}_{n=0}^{\infty}$ be a sequence of nonnegative numbers such that the power series $f(x) = \sum_{n=0}^{\infty} b_n x^n$ converges for $|x| < 1$. If*
$$\log b_n \sim 2\sqrt{\alpha n} \qquad \text{as } n \to \infty, \tag{16.4}$$
then
$$\log f(x) \sim \frac{\alpha}{1-x} \qquad \text{as } x \to 1^-. \tag{16.5}$$

Proof. Let $0 < \varepsilon < 1$. The asymptotic formula (16.4) implies that there exists a positive integer $N_0 = N_0(\varepsilon)$ such that

$$e^{2(1-\varepsilon)\sqrt{\alpha n}} < b_n < e^{2(1+\varepsilon)\sqrt{\alpha n}} \quad \text{for all } n \geq N_0.$$

The series $f(x)$ converges for $|x| < 1$ (by the root test), but diverges for $x = 1$. For $0 < x < 1$ we let $x = e^{-t}$, where $t = t(x) = -\log x > 0$, and t decreases to 0 as x increases to 1.

First, we derive the lower bound

$$\liminf_{x \to 1^-} (1 - x) \log f(x) \geq \alpha.$$

For $n \geq N_0$,

$$b_n x^n > e^{2(1-\varepsilon)\sqrt{\alpha n}} e^{-tn} = e^{2(1-\varepsilon)\sqrt{\alpha n} - tn}.$$

Completing the square in the exponent, we obtain

$$2(1-\varepsilon)\sqrt{\alpha n} - tn = \frac{(1-\varepsilon)^2 \alpha}{t} - t\left(\sqrt{n} - \frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2,$$

and so

$$b_n x^n > e^{\frac{(1-\varepsilon)^2 \alpha}{t}} e^{-t\left(\sqrt{n} - \frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2}.$$

Choose $t_0 > 0$ such that

$$\left(\frac{(1-\varepsilon)\sqrt{\alpha}}{t_0}\right)^2 > N_0 + 1,$$

and let $x_0 = e^{-t_0} \in (0, 1)$. Let $x_0 < x < 1$. If $x = e^{-t}$, then $0 < t < t_0$. Let

$$n_x = \left[\left(\frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2\right].$$

Then

$$N_0 < \left(\frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2 - 1 < n_x \leq \left(\frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2$$

and

$$\frac{(1-\varepsilon)\sqrt{\alpha}}{t} - 1 < \sqrt{\left(\frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2 - 1} < \sqrt{n_x} \leq \frac{(1-\varepsilon)\sqrt{\alpha}}{t}.$$

It follows that

$$\left(\sqrt{n_x} - \frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2 < 1,$$

and so

$$b_{n_x} x^{n_x} > e^{\frac{(1-\varepsilon)^2 \alpha^2}{t}} e^{-t\left(\sqrt{n_x} - \frac{(1-\varepsilon)\sqrt{\alpha}}{t}\right)^2} > e^{\frac{(1-\varepsilon)^2 \alpha^2}{t} - t}.$$

Since $b_n x^n \geq 0$ for all $n \geq 0$, we have

$$f(x) = \sum_{n=0}^{\infty} b_n x^n \geq b_{n_x} x^{n_x} > e^{\frac{(1-\varepsilon)^2 \alpha}{t} - t}.$$

Therefore,

$$\log f(x) > \frac{(1-\varepsilon)^2 \alpha}{t} - t$$

and

$$t \log f(x) > (1-\varepsilon)^2 \alpha - t^2.$$

By Exercise 1,

$$t = -\log x \sim 1 - x \quad \text{as } x \to 1^-,$$

and so

$$\begin{aligned} \liminf_{x \to 1^-} (1-x) \log f(x) &= \liminf_{x \to 1^-} t \log f(x) \\ &\geq \liminf_{t \to 0^+} \left((1-\varepsilon)^2 \alpha - t^2 \right) \\ &= (1-\varepsilon)^2 \alpha. \end{aligned}$$

This is true for $0 < \varepsilon < 1$, and so

$$\liminf_{x \to 1^-} (1-x) \log f(x) \geq \alpha.$$

Next we derive the upper bound

$$\limsup_{x \to 1^-} (1-x) \log f(x) \leq \alpha.$$

We have

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} b_n x^n \\ &< \sum_{n=0}^{N_0 - 1} b_n x^n + \sum_{n=N_0}^{\infty} e^{2(1+\varepsilon)\sqrt{\alpha n} - tn} \\ &\leq c_1(\varepsilon) + e^{\frac{(1+\varepsilon)^2 \alpha}{t}} \sum_{n=N_0}^{\infty} e^{-t\left(\sqrt{n} - \frac{(1+\varepsilon)\sqrt{\alpha}}{t}\right)^2}, \end{aligned}$$

where

$$0 \leq \sum_{n=0}^{N_0 - 1} b_n x^n \leq \sum_{n=0}^{N_0 - 1} b_n = c_1(\varepsilon).$$

Let

$$N_1 = N_1(t) = \left[\frac{16\alpha}{t^2} \right].$$

Then

$$\frac{4\alpha}{t} < \frac{t(N_1+1)}{4}.$$

If $n > N_1$, then

$$\sqrt{n} > \frac{4\sqrt{\alpha}}{t} > \frac{2(1+\varepsilon)\sqrt{\alpha}}{t}$$

and

$$\sqrt{n} - \frac{(1+\varepsilon)\sqrt{\alpha}}{t} > \frac{\sqrt{n}}{2}.$$

It follows that

$$e^{-t\left(\sqrt{n}-\frac{(1+\varepsilon)\sqrt{\alpha}}{t}\right)^2} < e^{-t\left(\frac{\sqrt{n}}{2}\right)^2} = e^{-\frac{tn}{4}},$$

and so, as $t \to 0^+$,

$$\begin{aligned}
\sum_{n=N_1+1}^{\infty} e^{-t\left(\sqrt{n}-\frac{(1+\varepsilon)\sqrt{\alpha}}{t}\right)^2} &< \sum_{n=N_1+1}^{\infty} e^{-tn/4} \\
&= \frac{e^{-t(N_1+1)/4}}{1-e^{-t/4}} \\
&< \frac{e^{-4\alpha/t}}{1-e^{-t/4}} \\
&< \frac{8e^{-4\alpha/t}}{t} \\
&= o(1),
\end{aligned}$$

since $1 - t/4 < e^{-t/4} < 1 - t/8$ for $0 < t < 1$. Also,

$$\sum_{n=N_0}^{N_1} e^{-t\left(\sqrt{n}-\frac{(1+\varepsilon)\sqrt{\alpha}}{t}\right)^2} < N_1 \leq \frac{16\alpha}{t^2}.$$

Consequently,

$$\begin{aligned}
f(x) &\leq c_1(\varepsilon) + e^{\frac{(1+\varepsilon)^2\alpha}{t}}\left(\frac{16\alpha}{t^2} + o(1)\right) \\
&\leq \frac{c_2(\varepsilon)e^{\frac{(1+\varepsilon)^2\alpha}{t}}}{t^2}.
\end{aligned}$$

Therefore,

$$\log f(x) \leq \frac{(1+\varepsilon)^2\alpha}{t} + \log\frac{c_2(\varepsilon)}{t^2}$$

and

$$t\log f(x) \leq (1+\varepsilon)^2\alpha + t\log\frac{c_2(\varepsilon)}{t^2}.$$

Then

$$\limsup_{x\to 1^-}(1-x)\log f(x) = \limsup_{t\to 0^+} t\log f(x) \leq (1+\varepsilon)^2\alpha.$$

This is true for every $\varepsilon > 0$, and so

$$\limsup_{x\to 1^-}(1-x)\log f(x) \leq \alpha.$$

This completes the proof.

Next we prove a Tauberian theorem about power series with real, nonnegative coefficients.

Theorem 16.4 *Let $B = \{b_n\}_{n=0}^{\infty}$ be a sequence of nonnegative real numbers. If the power series*

$$f(x) = \sum_{n=0}^{\infty} b_n x^n$$

converges for $|x| < 1$ and if

$$f(x) \sim \frac{1}{1-x} \quad \text{as } x \to 1^-,$$

then

$$\sum_{k=0}^{n} b_k \sim n.$$

Proof. We begin by showing that for every polynomial $p(x)$ we have

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n p(x^n) = \int_0^1 p(x)dx. \tag{16.6}$$

Since both sides are linear in $p(x)$, it suffices to prove this for $p(x) = x^k$. We have

$$\begin{aligned}
(1-x)\sum_{n=0}^{\infty} b_n x^n p(x^n) &= (1-x)\sum_{n=0}^{\infty} b_n x^n x^{kn} \\
&= \frac{1-x}{1-x^{k+1}}(1-x^{k+1})\sum_{n=0}^{\infty} b_n x^{(k+1)n} \\
&= \frac{1}{1+x+\cdots+x^k}(1-x^{k+1})\sum_{n=0}^{\infty} b_n (x^{k+1})^n,
\end{aligned}$$

and so

$$\lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n p(x^n)$$

$$
\begin{aligned}
&= \lim_{x\to 1^-} \frac{1}{1+x+\cdots+x^k} \lim_{x\to 1^-} (1-x^{k+1}) \sum_{n=0}^{\infty} b_n \left(x^{k+1}\right)^n \\
&= \frac{1}{k+1} = \int_0^1 x^k dx.
\end{aligned}
$$

This proves (16.6).

Next we use the Weierstrass approximation theorem: If $f(x)$ is a continuous function on the interval $[0,1]$ and if $\varepsilon > 0$, then there exists a polynomial $p(x)$ such that

$$|f(x) - p(x)| < \varepsilon \quad \text{for all } x \in [0,1].$$

Let $f^+(x) = f(x) + \varepsilon/2$, and let $p^+(x)$ be a polynomial such that

$$|f^+(x) - p^+(x)| < \varepsilon/2 \quad \text{for all } x \in [0,1].$$

Then

$$f(x) < p^+(x) < f(x) + \varepsilon \quad \text{for all } x \in [0,1]$$

and

$$\int_0^1 f(x)dx < \int_0^1 p^+(x)dx < \int_0^1 f(x)dx + \varepsilon.$$

Similarly, there exists a polynomial $p^-(x)$ such that

$$f(x) - \varepsilon < p^-(x) < f(x) \quad \text{for all } x \in [0,1]$$

and

$$\int_0^1 f(x)dx - \varepsilon < \int_0^1 p^-(x)dx < \int_0^1 f(x)dx.$$

Consider the function

$$g(x) = \begin{cases} 0 & \text{for } 0 \le x < e^{-1}, \\ \frac{1}{x} & \text{for } e^{-1} \le x \le 1. \end{cases}$$

Then

$$\int_0^1 g(x)dx = \int_{e^{-1}}^1 \frac{dx}{x} = 1.$$

The function $g(x)$ is continuous for all $x \in [0,1]$ except for $x = e^{-1}$, where it has a jump discontinuity, and so we cannot apply Weierstrass's theorem directly to approximate $g(x)$ from above and below by polynomials. We circumvent this difficulty in the following way. Let $0 < \varepsilon < e^{-1}$. Define the function $f^+(x)$ as follows:

$$f^+(x) = \begin{cases} \frac{\varepsilon}{2} & \text{for } 0 \le x \le e^{-1} - \varepsilon, \\ \ell^+(x) & \text{for } e^{-1} - \varepsilon \le x \le e^{-1}, \\ \frac{1}{x} + \frac{\varepsilon}{2} & \text{for } e^{-1} \le x \le 1, \end{cases}$$

where $\ell^+(x)$ is the straight line with end points $(e^{-1}-\varepsilon, \varepsilon/2)$ and $(e^{-1}, e+\varepsilon/2)$. Then $f^+(x)$ is a continuous function on the interval $[0,1]$, and so there exists a polynomial $p^+(x)$ such that

$$g(x) < f^+(x) < p^+(x) < f^+(x) + \frac{\varepsilon}{2}$$

for all $x \in [0,1]$. Then

$$0 < p^+(x) < \begin{cases} \varepsilon & \text{for } 0 \leq x \leq e^{-1} - \varepsilon, \\ e + \varepsilon & \text{for } e^{-1} - \varepsilon \leq x \leq e^{-1}, \\ \frac{1}{x} + \varepsilon & \text{for } e^{-1} \leq x \leq 1, \end{cases}$$

and so

$$\begin{aligned} 1 &= \int_0^1 g(x)dx \\ &< \int_0^1 p^+(x)dx \\ &= \int_0^{e^{-1}-\varepsilon} p^+(x)dx + \int_{e^{-1}-\varepsilon}^{e^{-1}} p^+(x)dx + \int_{e^{-1}}^1 p^+(x)dx \\ &< \varepsilon(e^{-1}-\varepsilon) + (e+\varepsilon)\varepsilon + 1 + \varepsilon(1-e^{-1}) \\ &= 1 + (e+1)\varepsilon. \end{aligned}$$

Similarly, we define the function $f^-(x)$ as follows:

$$f^-(x) = \begin{cases} \frac{-\varepsilon}{2} & \text{for } 0 \leq x \leq e^{-1} \\ \ell^-(x) & \text{for } e^{-1} \leq x \leq e^{-1} + \varepsilon \\ \frac{1}{x} - \frac{\varepsilon}{2} & \text{for } e^{-1} + \varepsilon \leq x \leq 1, \end{cases}$$

where $\ell^-(x)$ is the straight line with end points $(e^{-1}, -\varepsilon/2)$ and $(e^{-1}+\varepsilon, 1/(e^{-1}-\varepsilon/2)$. Then $f^-(x)$ is a continuous function on the interval $[0,1]$, and there exists a polynomial $p^-(x)$ such that

$$f^-(x) - \frac{\varepsilon}{2} < p^-(x) < f^-(x) < g(x)$$

for all $x \in [0,1]$. It follows that

$$\begin{aligned} 1 &= \int_0^1 g(x)dx \\ &> \int_0^1 p^-(x)dx \\ &> \int_0^{e^{-1}+\varepsilon} (-\varepsilon)dx + \int_{e^{-1}+\varepsilon}^1 \left(\frac{1}{x} - \varepsilon\right) dx \\ &= -\varepsilon(e^{-1}+\varepsilon) - \log(e^{-1}+\varepsilon) - \varepsilon(1 - e^{-1} - \varepsilon) \\ &= 1 - \varepsilon - \log(1 + e\varepsilon) \\ &> 1 - (e+1)\varepsilon. \end{aligned}$$

The inequality $p^-(x) < g(x) < p^+(x)$ implies that for $0 < x < 1$,

$$\begin{aligned} (1-x)\sum_{n=0}^{\infty} b_n x^n p^-(x^n) &< (1-x)\sum_{n=0}^{\infty} b_n x^n g(x^n) \\ &< (1-x)\sum_{n=0}^{\infty} b_n x^n p^+(x^n). \end{aligned}$$

By (16.6),

$$\begin{aligned} 1-(e+1)\varepsilon &< \int_0^1 p^-(t)dt \\ &= \lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n p^-(x^n) \\ &\leq \liminf_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n g(x^n) \\ &\leq \limsup_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n g(x^n) \\ &\leq \lim_{x\to 1^-}(1-x)\sum_{n=0}^{\infty} b_n x^n p^+(x^n) \\ &= \int_0^1 p^+(x)dx \\ &< 1+(e+1)\varepsilon. \end{aligned}$$

These inequalities hold for all sufficiently small ε, and so

$$\lim_{x\to 1^-}(1-x)\sum_{k=0}^{\infty} b_k x^k g(x^k) = 1.$$

Let

$$x = e^{-1/n}.$$

Then $0 < x < 1$, and

$$e^{-1} \leq x^k = e^{-k/n} \leq 1$$

if and only if

$$k = 0, 1, \ldots, n.$$

It follows from the definition of the function $g(x)$ that

$$\sum_{k=0}^{\infty} b_k x^k g(x^k) = \sum_{k=0}^{n} b_k x^k g(x^k) = \sum_{k=0}^{n} b_k,$$

and so

$$\lim_{n\to\infty}(1-e^{-1/n})\sum_{k=0}^{n} b_k = 1,$$

that is,

$$\sum_{k=0}^{n} b_k \sim \frac{1}{1-e^{-1/n}}.$$

From the inequality

$$1-x < e^{-x} < 1-x+\frac{x^2}{2}$$

with $x = 1/n$, we obtain

$$\frac{1}{n}\left(1-\frac{1}{2n}\right) < 1-e^{-1/n} < \frac{1}{n},$$

and so

$$\frac{1}{1-e^{-1/n}} \sim n$$

as $n \to \infty$. Therefore,

$$\sum_{k=0}^{n} b_k \sim n.$$

This completes the proof.

Exercises

1. Prove that
$$-\log x \sim 1-x \quad \text{as } x \to 1^-.$$

2. Let $B = \{b_n\}_{n=0}^{\infty}$ be a sequence of real, nonnegative numbers such that the power series $f(x) = \sum_{n=0}^{\infty} b_n x^n$ converges for $|x| < 1$. Prove that if
$$\liminf_{n\to\infty} \frac{\log b_n}{2\sqrt{n}} \geq \sqrt{\alpha},$$
then
$$\liminf_{x\to 1^-}(1-x)\log f(x) \geq \alpha.$$

3. Let $B = \{b_n\}_{n=0}^{\infty}$ be a sequence of real, nonnegative numbers such that the power series $f(x) = \sum_{n=0}^{\infty} b_n x^n$ converges for $|x| < 1$. Prove that if
$$\limsup_{n\to\infty} \frac{\log b_n}{2\sqrt{n}} \leq \sqrt{\alpha},$$
then
$$\limsup_{x\to 1^-}(1-x)\log f(x) \leq \alpha.$$

16.4 Notes

Theorem 16.1 and Theorem 16.2 show that a set A with $\gcd(A) = 1$ has positive density α if and only if $\log p_A(n) \sim c_0\sqrt{\alpha n}$. Erdős states these results, with a sketch of a proof, in his paper [32], where Theorem 16.3 is also stated and applied. The proofs in this book appear in Nathanson [105, 106].

Theorem 16.4 is a famous Tauberian theorem of Hardy and Littlewood [53]; the proof in this book is due to Karamata [77]. Titchmarsh [142, Chapter 7] discusses this and many related results.

Using hard analytic machinery, Freiman [36], Kohlbecker [84], and Yang [158] have obtained other inverse theorems for partitions.

We know the asymptotics of partition functions for certain sets of integers of zero density. For example, Hardy and Ramanujan [57] proved that if $A^{(k)}$ is the set of kth powers of positive integers, then

$$\log p_{A^{(k)}}(n) \sim (k+1)\left\{\frac{1}{k}\Gamma\left(\frac{1}{k}+1\right)\zeta\left(\frac{1}{k}+1\right)\right\}^{k/(k+1)} n^{1/(k+1)},$$

where $\Gamma(s)$ is the gamma function and $\zeta(s)$ is the Riemann zeta function. This gives (15.2) in the special case $k = 1$. In the same paper, they also proved that if $\mathbf{P}$ is the set of prime numbers, then

$$\log p_{\mathbf{P}}(n) \sim 2\pi\sqrt{\frac{n}{3\log n}},$$

and if $\mathbf{P}^{(k)}$ is the set of kth powers of primes, then

$$\log p_{\mathbf{P}^{(k)}}(n) \sim (k+1)\left\{\Gamma\left(\frac{1}{k}+2\right)\zeta\left(\frac{1}{k}+1\right)\right\}^{k/(k+1)}\left\{\frac{n}{(\log n)^k}\right\}^{1/(k+1)}.$$

References

[1] W. R. Alford, A. Granville, and C. Pomerance. There are infinitely many Carmichael numbers. *Annals Math.*, 139:703–722, 1994.

[2] N. Alon, M. B. Nathanson, and I. Ruzsa. The polynomial method and restricted sums of congruence classes. *J. Number Theory*, 56:404–417, 1996.

[3] T. M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1976.

[4] T. M. Apostol. *Modular Forms and Dirichlet Series in Number Theory*, volume 41 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1989.

[5] E. Artin. *Collected Papers*. Springer-Verlag, New York, 1965.

[6] F. C. Auluck, S. Chowla, and H. Gupta. On the maximum value of the number of partitions of n into k parts. *J. Indian Math. Soc. (N.S.)*, 6:105–112, 1942.

[7] L. Auslander and R. Tolimieri. Ring structure and the Fourier transform. *Math. Intelligencer*, 7(3):49–52, 54, 1985.

[8] B. C. Berndt and R. J. Evans. The determination of Gauss sums. *Bull. Amer. Math. Soc.*, 5:107–129, 1981.

[9] B. C. Berndt, R. J. Evans, and K. S. Williams. *Gauss and Jacobi Sums*. John Wiley & Sons, New York, 1998.

[10] A. S. Besicovitch. On the density of the sum of two sequences of integers. *Math. Annalen*, 110:336–341, 1934.

[11] H. Bohr. Address of Professor Harald Bohr. In *Proceedings of the International Congress of Mathematicians (Cambridge, 1950)*, volume 1, pages 127–134, Providence, 1952. Amer. Math. Soc.

[12] D. Boneh. Twenty years of attacks on the RSA cryptosystem. *Notices Amer. Math. Soc.*, 46:203–213, 1999.

[13] Z. I. Borevich and I. R. Shafarevich. *Number Theory.* Academic Press, New York, 1966.

[14] J. Browkin and J. Brzeziński. Some remarks on the *abc*-conjecture. *Math. Comp.*, 62:931–939, 1994.

[15] J. Brzeziński. The *abc*-conjecture. Preprint, 1999.

[16] S. Chowla. On abundant numbers. *J. Indian Math. Soc. (2)*, 1:41–44, 1934.

[17] R. Crandall, K. Dilcher, and C. Pomerance. A search for Wieferich and Wilson primes. *Math. Comp.*, 66:433–449, 1997.

[18] H. Daboussi. Sur le théorème des nombres premiers. *Comptes Rendus Acad. Sci. Paris, Sér. A*, 298:161–164, 1984.

[19] H. Davenport. Über numeri abundantes. *Sitzungsbericht Aka. Wiss. Berlin*, 27:830–837, 1933.

[20] H. Davenport. On $f^3(t) - g^2(t)$. *Norske Vid. Selsk. Forrh.*, 38:86–87, 1965.

[21] H. Davenport. *Multiplicative Number Theory*, volume 74 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2nd edition, 1980.

[22] H. Davenport. *The Higher Arithmetic.* Cambridge University Press, Cambridge, 6th edition, 1992.

[23] C.-J. de la Vallée Poussin. Recherches analytiques sur la théorie des nombres; Première partie: La function $\zeta(s)$ de Riemann et les nombres premiers en général. *Annales de la Soc. scientifique de Bruxelles*, 20:183–256, 1896.

[24] H. G. Diamond. Elementary methods in the study of the distribution of prime numbers. *Bull. Am. Math. Soc.*, 7:553–589, 1982.

[25] L. E. Dickson. *History of the Theory of Numbers.* Carnegie Institute of Washington, Washington, 1919, 1920, 1923; reprinted by Chelsea Publishing Company in 1971.

[26] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, IT–22:644–654, 1976.

[27] J. S. Ellenberg. Congruence ABC implies ABC. Preprint, 1999.

[28] P. T. D. A. Elliott. *Probabilistic Number Theory I: Mean Value Theorems.* Springer-Verlag, New York, 1979.

[29] P. T. D. A. Elliott. *Probabilistic Number Theory II: Central Limit Theorems.* Springer-Verlag, New York, 1980.

[30] P. T. D. A. Elliott. The multiplicative group of rationals generated by the shifted primes, I. *J. reine angew. Math.*, 463:169–216, 1995.

[31] P. Erdős. On the density of the abundant numbers. *J. London Math. Soc.*, 9:278–282, 1934.

[32] P. Erdős. On an elementary proof of some asymptotic formulas in the theory of partitions. *Annals Math.*, 43:437–450, 1942.

[33] P. Erdős. On some asymptotic formulas in the theory of partitions. *Bull. Amer. Math. Soc.*, 52:185–188, 1946.

[34] P. Erdős. On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. *Proc. Nat. Acad. Sci. U.S.A.*, 35:374–384, 1949.

[35] P. Erdős and J. Lehner. The distribution of the number of summands in the partitions of a positive integer. *Duke Math. J.*, 8:335–345, 1941.

[36] G. A. Freiman. Inverse problems of the additive theory of numbers. *Izv. Akad. Nauk SSSR*, 19:275–284, 1955.

[37] C. F. Gauss. *Disquisitiones Arithmeticae.* Springer-Verlag, New York, 1986. Translated by A. A. Clarke and revised by W. C. Waterhouse.

[38] D. Goldfeld. The elementary proof of the prime number theorem: An historical perspective. Preprint, 1998.

[39] A. Granville. On elementary proofs of the Prime Number Theorem for arithmetic progressions, without characters. In *Proceedings of the Amalfi Conference on Analytic Number Theory, September 25–29, 1989*, pages 157–195, Salerno, Italy, 1992. Universitá di Salerno.

[40] A. Granville. Primality testing and Carmichael testing. *Notices Amer. Math. Soc.*, 39:696–700, 1992.

[41] N. Greenleaf. On Fermat's equation in $\mathbf{C}(t)$. *Am. Math. Monthly*, 76:808–809, 1969.

[42] E. Grosswald. *Topics from the Theory of Numbers.* Macmillan, New York, 1966.

[43] E. Grosswald. *Representations of Integers as Sums of Squares.* Springer-Verlag, New York, 1985.

[44] R. Gupta and M. R. Murty. A remark on Artin's conjecture. *Inventiones Math.*, 78:127–130, 1984.

[45] R. K. Guy. *Unsolved Problems in Number Theory.* Springer-Verlag, New York, 2 edition, 1994.

[46] J. Hadamard. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bulletin de la Soc. math. de France*, 24:199–220, 1896.

[47] H. Halberstam and H.-E. Richert. *Sieve Methods.* Academic Press, London, 1974.

[48] H. Halberstam and K. F. Roth. *Sequences*, volume 1. Oxford University Press, Oxford, 1966. Reprinted by Springer-Verlag, Heidelberg, in 1983.

[49] R. R. Hall. *Sets of Multiples.* Number 118 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996.

[50] R. R. Hall and G. Tenenbaum. *Divisors.* Number 90 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1988.

[51] G. H. Hardy. *A Mathematician's Apology.* Cambridge University Press, Cambridge, 1940. Reprinted in 1967.

[52] G. H. Hardy. *Ramanujan. Twelve Lectures on Subjects Suggested by his Life and Work.* Cambridge University Press, Cambridge, 1940. Reprinted by Chelsea Publishing Company, New York, in 1959.

[53] G. H. Hardy and J. E. Littlewood. Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive. *Proc. London Math. Soc.*, 13:174–191, 1914.

[54] G. H. Hardy and J. E. Littlewood. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. *Acta Math.*, 41:119–196, 1918.

[55] G. H. Hardy and J. E. Littlewood. A new solution of Waring's problem. *Q. J. Math.*, 48:272–293, 1919.

[56] G. H. Hardy and J. E. Littlewood. Some problems of "Partitio Numerorum." A new solution of Waring's problem. *Göttingen Nach.*, pages 33–54, 1920.

[57] G. H. Hardy and S. Ramanujan. Asymptotic formulae for the distribution of integers of various types. *Proc. London Math. Soc.*, 16:112–132, 1917.

[58] G. H. Hardy and S. Ramanujan. Asymptotic formulae in combinatory analysis. *Proc. London Math. Soc.*, 17:75–115, 1918.

[59] G. H. Hardy and S. Ramanujan. Une formule asymptotique pour le nombres des partitions de n. *Comptes Rendus Acad. Sci. Paris, Sér. A*, 2 Jan. 1917.

[60] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, Oxford, 5th edition, 1979.

[61] T. L. Heath. *The Thirteen Books of Euclid's Elements*. Dover Publications, New York, 1956.

[62] D. R. Heath-Brown. Artin's conjecture for primitive roots. *Quart. J. Math. Oxford*, 37:22–38, 1986.

[63] E. Hecke. *Vorlesungen über die Theorie der Algebraischen Zahlen*. Akademische Verlagsgesellschaft, Leipzig, 1923. Reprinted by Chelsea Publishing Company, New York, in 1970.

[64] E. Hecke. *Lectures on the Theory of Algebraic Numbers*, volume 77 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1981.

[65] M. E. Hellman. The mathematics of public-key cryptography. *Scientific American*, 241:130–139, 1979.

[66] D. Hilbert. Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem). *Mat. Annalen*, 67:281–300, 1909.

[67] A. Hildebrand. The Prime Number Theorem via the large sieve. *Mathematika*, 33:23–30, 1986.

[68] L. K. Hua. *Introduction to Number Theory*. Springer-Verlag, Berlin, 1982.

[69] A. E. Ingham. Some asymptotic formulae in the theory of numbers. *J. London Math. Soc.*, 2:202–208, 1927.

[70] A. E. Ingham. *The Distribution of Prime Numbers*. Number 30 in Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge University Press, Cambridge, 1932. Reprinted in 1992.

[71] A. E. Ingham. Review of the papers of Selberg and Erdős. *Math. Reviews*, 10(595b, 595c), 1949. Reprinted in [92, vol. 4, pages 191–193, N20–3].

[72] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1990.

[73] H. Iwaniec. Almost-primes represented by quadratic polynomials. *Inventiones Math.*, 47:171–188, 1978.

[74] H. Iwaniec. *Topics in Classical Automorphic Forms*, volume 17 of *Graduate Studies in Mathematics*. Amer. Math. Soc., Providence, 1997.

[75] S. M. Johnson. On the representations of an integer as a sum of products. *Trans. Amer. Math. Soc.*, 76:177–189, 1954.

[76] E. Kamke. Verallgemeinerung des Waring-Hilbertschen Satzes. *Math. Annalen*, 83:85–112, 1921.

[77] J. Karamata. Über die Hardy–Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes. *Math. Zeit.*, 32:319–320, 1930.

[78] A. Ya. Khinchin. *Three Pearls of Number Theory*. Dover Publications, Mineola, NY, 1998. This translation from the Russian of the second (1948), revised edition was published originally by Graylock Press in 1952.

[79] M. Kneser. Abschätzungen der asymptotischen Dichte von Summenmengen. *Math. Z.*, 58:459–484, 1953.

[80] C. Knessl and J. B. Keller. Partition asymptotics for recursion equations. *SIAM J. Applied Math.*, 50:323–338, 1990.

[81] M. I. Knopp. *Modular Functions in Analytic Number Theory*. Markham Publishing Co., Chicago, 1970. Reprinted by Chelsea Publishing Company in 1993.

[82] Chao Ko. On the diophantine equation $x^2 = y^n + 1, xy \neq 0$. *Scientia Sinica*, 14:457–460, 1964.

[83] N. Koblitz. *A Course in Number Theory and Cryptography*, volume 114 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1994.

[84] E. E. Kohlbecker. Weak asymptotic properties of partitions. *Trans. Amer. Math. Soc.*, 88:346–375, 1958.

[85] R. Kumanduri and C. Romero. *Number Theory with Computer Applications*. Prentice Hall, Upper Saddle River, New Jersey, 1998.

[86] A. V. Kuzel'. Elementary solution of Waring's problem for polynomials by the method of Yu. B. Linnik. *Uspekhi Mat. Nauk*, 11:165–168, 1956.

[87] E. Landau. *Elementary Number Theory*. Chelsea Publishing Company, New York, 1966.

[88] S. Lang. Old and new conjectured diophantine inequalities. *Bull. Amer. Math. Soc.*, 23:37–75, 1990.

[89] S. Lang. *Algebra*. Addison-Wesley, Reading, Mass., 3rd edition, 1993.

[90] S. Lang. *Algebraic Number Theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1994.

[91] V. A. Lebesgue. Sur l'impossibilité, en nombres entiers, de l'équation $x^m = y^2 + 1$. *Nouv. Ann. Math. (1)*, 9:178–181, 1850.

[92] W. J. LeVeque. *Reviews in Number Theory*. Amer. Math. Soc., Providence, 1974.

[93] Yu. V. Linnik. An elementary solution of Waring's problem by Shnirel'man's method. *Mat. Sbornik NS*, 12 (54):225–230, 1943.

[94] J. E. Littlewood. Sur la distribution des nombres premiers. *C. R. Acad. Sci. Paris, Sér. A*, 158:1869–1872, 1914.

[95] Yu. I. Manin. Classical computing, quantum computing, and Shor's factorization algorithm. In *Séminaire Bourbaki, 51ème année, 1998–99*, pages 862–1—862–30. UFR de Mathématiques de l'Université Paris VII — Denis Diderot, Paris, 1999.

[96] Yu. I. Manin and A. A. Panchishkin. *Number Theory I. Introduction to Number Theory*, volume 49 of *Encyclopedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1995.

[97] R. C. Mason. *Diophantine Equations over Function Fields*, volume 96 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, Cambridge, 1984.

[98] M. R. Murty. Artin's conjecture for primitive roots. *Math. Intelligencer*, 10(4):59–67, 1988.

[99] A. P. Nathanson. "Arithmetic". Poem written in D'Ann Ippolito's third grade class at Far Brook School, 1998.

[100] M. B. Nathanson. An exponential congruence of Mahler. *Amer. Math. Monthly*, 79:55–57, 1972.

[101] M. B. Nathanson. Sums of finite sets of integers. *Amer. Math. Monthly*, 79:1010–1012, 1972.

[102] M. B. Nathanson. Catalan's equation in $K(t)$. *Amer. Math. Monthly*, 81:371–373, 1974.

[103] M. B. Nathanson. *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, volume 165 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[104] M. B. Nathanson. *Additive Number Theory: The Classical Bases*, volume 164 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[105] M. B. Nathanson. On Erdős's elementary method in the asymptotic theory of partitions. Preprint, 1998.

[106] M. B. Nathanson. Asymptotic density and the asymptotics of partition functions. *Acta Math. Hungar.*, 87(1–2), 2000.

[107] M. B. Nathanson. Partitions with parts in a finite set. *Proc. Amer. Math. Soc.*, 2000. To appear.

[108] M. B. Nathanson. Additive Number Theory: Addition Theorems and the Growth of Sumsets. In preparation, 2001.

[109] V. I. Nechaev. *Waring's Problem for Polynomials*. Izdat. Akad. Nauk SSSR, Moscow, 1951.

[110] O. Neugebauer. *The Exact Sciences in Antiquity*. Brown Univ. Press, Providence, 2nd edition, 1957. Reprinted by Dover Publications in 1969.

[111] J. Neukirch. *Algebraic Number Theory*. Springer-Verlag, Berlin, 1999.

[112] D. J. Newman. Simple analytic proof of the prime number theorem. *Amer. Math. Monthly*, 87:693–696, 1980.

[113] A. Nitaj. La conjecture *abc*. *Enseignement Math.*, 42:3–24, 1996.

[114] J. Oesterlé. Nouvelles approches du "Théorème" de Fermat. In *Séminaire Bourbaki, Volume 1987/88, Exposés 686–699*, volume 161–162 of *Astérisque*. Société Mathématique de France, Paris, 1988.

[115] A. G. Postnikov. A remark on an article by A. G. Postnikov and N. P. Romanov. *Uspehki Mat. Nauk*, 24(5(149)):263, 1969.

[116] A. G. Postnikov and N. P. Romanov. A simplification of A. Selberg's elementary proof of the asymptotic law of distribution of prime numbers. *Uspehki Mat. Nauk (N.S.)*, 10(4(66)):75–87, 1955.

[117] H. Rademacher. A convergent series for the partition function $p(n)$. *Proc. Nat. Acad. Sci.*, 23:78–84, 1937.

[118] H. Rademacher. On the partition function $p(n)$. *Proc. London Math. Soc.*, 43:241–254, 1937.

[119] H. Rademacher. *Topics in Analytic Number Theory.* Springer-Verlag, New York, 1973.

[120] D. Ramakrishnan and R. J. Valenza. *Fourier Analysis on Number Fields*, volume 186 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1999.

[121] S. Ramanujan. Some formulæ in the analytic theory of numbers. *Messenger of Mathemtics*, 45:81–84, 1916.

[122] G. J. Rieger. Zu Linniks Lösung des Waringschen Problems: Abschätzung von $g(n)$. *Math. Zeit.*, 60:213–239, 1954.

[123] R. L. Rivest, A. Shamir, and L. M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21:120–126, 1978.

[124] A. Schinzel. Remarks on the paper "Sur certaines hypothèses concernant les nombres premiers". *Acta Arith.*, 7:1–8, 1961/62.

[125] A. Schinzel and W. Sierpiński. Sur certaines hypothèses concernant les nombres premiers. *Acta Arith.*, 4:185–208, 1958. Erratum 5 (1959), 259.

[126] I. Schur. Über die Gaußschen Summen. *Nachrichten k. Gesell. Göttingen, Math.-Phys. Klasse*, pages 147–153, 1921. Reprinted in *Gesammelte Abhandlungen*, Band II, Springer-Verlag, Berlin, 1973.

[127] A. Selberg. An elementary proof of Dirichlet's theorem about primes in an arithmetic progression. *Annals Math.*, 50:297–304, 1949. In *Collected Papers*, volume I, pages 371–378, Springer-Verlag, Berlin, 1989.

[128] A. Selberg. An elementary proof of the prime-number theorem. *Annals Math.*, 50:305–313, 1949. In *Collected Papers*, volume I, pages 379–387, Springer-Verlag, Berlin, 1989.

[129] A. Selberg. An elementary proof of the prime-number theorem for arithmetic progressions. *Canadian J. Math.*, 2:66–78, 1950. In *Collected Papers*, volume I, pages 398–410, Springer-Verlag, Berlin, 1989.

[130] A. Selberg. Reflections around the Ramanujan centenary. In *Collected Papers*, volume I, pages 695–706. Springer-Verlag, Berlin, 1989.

[131] J.-P. Serre. *Cours d'Arithmétique.* Presses Universitaires de France, Paris, 1970.

[132] J.-P. Serre. *A Course in Arithmetic*, volume 7 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1973.

[133] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26:1484–1509, 1997.

[134] J. H. Silverman. Wieferich's criterion and the *abc* conjecture. *J. Number Theory*, 30:226–237, 1988.

[135] S. Singh. *The Code Book: The Evolution of Secrecy from Mary, Queen of Scots to Quantum Cryptography.* Doubleday, New York, 1999.

[136] E. G. Straus. The elementary proof of the Prime Number Theorem. Undated, unpublished manuscript.

[137] G. Szekeres. An asymptotic formula in the theory of partitions. *Quarterly. J. Math. Oxford*, 2:85–108, 1951.

[138] G. Szekeres. Some asymptotic formulae in the theory of partitions (II). *Quarterly. J. Math. Oxford*, 4:96–111, 1953.

[139] R. Taylor and A. Wiles. Ring-theoretic properties of certain Hecke algebras. *Annals Math.*, 141:533–572, 1995.

[140] G. Tenenbaum and M. Mendès-France. *The Prime Numbers and Their Distribution.* Amer. Math. Soc., Providence, 1999.

[141] A. Terras. *Fourier Analysis on Finite Groups and Applications.* Number 43 in London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1999.

[142] E. C. Titchmarsh. *The Theory of Functions.* Oxford University Press, Oxford, 2nd edition, 1939.

[143] P. Turán. On a theorem of Hardy and Ramanujan. *J. London Math. Soc.*, 9:274–276, 1934.

[144] P. Turán. *On a New Method of Analysis and its Applications.* Wiley-Interscience, New York, 1984.

[145] J. V. Uspensky and M. A. Heaslet. *Elementary Number Theory.* McGraw-Hill, New York, 1939.

[146] Ya. V. Uspensky. Asymptotic expressions of numerical functions occurring in problems concerning the partition of numbers into summands. *Bull. Acad. Sci. de Russie*, 14(6):199–218, 1920.

[147] B. L. van der Waerden. *Science Awakening.* Science Editions, John Wiley & Sons, New York, 2nd edition, 1963.

[148] R. C. Vaughan. *The Hardy–Littlewood Method.* Cambridge University Press, Cambridge, 2nd edition, 1997.

[149] B. A. Venkov. *Elementary Number Theory.* Wolters-Noordhof Publishing, Groningen, the Netherlands, 1970.

[150] I. M. Vinogradov. On Waring's theorem. *Izv. Akad. Nauk SSSR, Otd. Fiz.-Mat. Nauk*, (4):393–400, 1928. English translation in *Selected Works*, pages 101–106, Springer-Verlag, Berlin, 1985.

[151] S. S. Wagstaff. Solution of Nathanson's exponential congruence. *Math. Comp.*, 33:1097–1100, 1979.

[152] A. Weil. *Number Theory for Beginners.* Springer-Verlag, New York, 1979.

[153] A. Weil. *Number Theory: An Approach through History. From Hammurapi to Legendre.* Birkhäuser, Boston, 1984.

[154] A. Weil. *Basic Number Theory.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 3rd edition, published in 1974.

[155] A. Wieferich. Zum letzten Fermat'schen Satz. *J. reine angew. Math.*, 136:293–302, 1909.

[156] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Annals Math.*, 141:443–531, 1995.

[157] B. M. Wilson. Proofs of some formulæ enunciated by Ramanujan. *Proc. London Math. Soc.*, 21:235–255, 1922.

[158] Y. Yang. Inverse problems for partition functions. Preprint, 1998.

[159] D. Zagier. Newman's short proof of the prime number theorem. *Amer. Math. Monthly*, 104:705–708, 1997.

Index

Graduate Texts in Mathematics

(continued from page ii)

120 ZIEMER. Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.
121 LANG. Cyclotomic Fields I and II. Combined 2nd ed.
122 REMMERT. Theory of Complex Functions. *Readings in Mathematics*
123 EBBINGHAUS/HERMES et al. Numbers. *Readings in Mathematics*
124 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part III.
125 BERENSTEIN/GAY. Complex Variables: An Introduction.
126 BOREL. Linear Algebraic Groups. 2nd ed.
127 MASSEY. A Basic Course in Algebraic Topology.
128 RAUCH. Partial Differential Equations.
129 FULTON/HARRIS. Representation Theory: A First Course. *Readings in Mathematics*
130 DODSON/POSTON. Tensor Geometry.
131 LAM. A First Course in Noncommutative Rings.
132 BEARDON. Iteration of Rational Functions.
133 HARRIS. Algebraic Geometry: A First Course.
134 ROMAN. Coding and Information Theory.
135 ROMAN. Advanced Linear Algebra.
136 ADKINS/WEINTRAUB. Algebra: An Approach via Module Theory.
137 AXLER/BOURDON/RAMEY. Harmonic Function Theory.
138 COHEN. A Course in Computational Algebraic Number Theory.
139 BREDON. Topology and Geometry.
140 AUBIN. Optima and Equilibria. An Introduction to Nonlinear Analysis.
141 BECKER/WEISPFENNING/KREDEL. Gröbner Bases. A Computational Approach to Commutative Algebra.
142 LANG. Real and Functional Analysis. 3rd ed.
143 DOOB. Measure Theory.
144 DENNIS/FARB. Noncommutative Algebra.
145 VICK. Homology Theory. An Introduction to Algebraic Topology. 2nd ed.
146 BRIDGES. Computability: A Mathematical Sketchbook.
147 ROSENBERG. Algebraic *K*-Theory and Its Applications.
148 ROTMAN. An Introduction to the Theory of Groups. 4th ed.
149 RATCLIFFE. Foundations of Hyperbolic Manifolds.
150 EISENBUD. Commutative Algebra with a View Toward Algebraic Geometry.
151 SILVERMAN. Advanced Topics in the Arithmetic of Elliptic Curves.
152 ZIEGLER. Lectures on Polytopes.
153 FULTON. Algebraic Topology: A First Course.
154 BROWN/PEARCY. An Introduction to Analysis.
155 KASSEL. Quantum Groups.
156 KECHRIS. Classical Descriptive Set Theory.
157 MALLIAVIN. Integration and Probability.
158 ROMAN. Field Theory.
159 CONWAY. Functions of One Complex Variable II.
160 LANG. Differential and Riemannian Manifolds.
161 BORWEIN/ERDÉLYI. Polynomials and Polynomial Inequalities.
162 ALPERIN/BELL. Groups and Representations.
163 DIXON/MORTIMER. Permutation Groups.
164 NATHANSON. Additive Number Theory: The Classical Bases.
165 NATHANSON. Additive Number Theory: Inverse Problems and the Geometry of Sumsets.
166 SHARPE. Differential Geometry: Cartan's Generalization of Klein's Erlangen Program.
167 MORANDI. Field and Galois Theory.
168 EWALD. Combinatorial Convexity and Algebraic Geometry.
169 BHATIA. Matrix Analysis.
170 BREDON. Sheaf Theory. 2nd ed.
171 PETERSEN. Riemannian Geometry.
172 REMMERT. Classical Topics in Complex Function Theory.
173 DIESTEL. Graph Theory.
174 BRIDGES. Foundations of Real and Abstract Analysis.
175 LICKORISH. An Introduction to Knot Theory.
176 LEE. Riemannian Manifolds.
177 NEWMAN. Analytic Number Theory.
178 CLARKE/LEDYAEV/STERN/WOLENSKI. Nonsmooth Analysis and Control Theory.
179 DOUGLAS. Banach Algebra Techniques in Operator Theory. 2nd ed.

180 SRIVASTAVA. A Course on Borel Sets.
181 KRESS. Numerical Analysis.
182 WALTER. Ordinary Differential Equations.
183 MEGGINSON. An Introduction to Banach Space Theory.
184 BOLLOBAS. Modern Graph Theory.
185 COX/LITTLE/O'SHEA. Using Algebraic Geometry.
186 RAMAKRISHNAN/VALENZA. Fourier Analysis on Number Fields.
187 HARRIS/MORRISON. Moduli of Curves.
188 GOLDBLATT. Lectures on the Hyperreals: An Introduction to Nonstandard Analysis.
189 LAM. Lectures on Modules and Rings.
190 ESMONDE/MURTY. Problems in Algebraic Number Theory.
191 LANG. Fundamentals of Differential Geometry.
192 HIRSCH/LACOMBE. Elements of Functional Analysis.
193 COHEN. Advanced Topics in Computational Number Theory.
194 ENGEL/NAGEL. One-Parameter Semigroups for Linear Evolution Equations.
195 NATHANSON. Elementary Methods in Number Theory.